arXiv:1110.1785v1 [cs.DM] 9 Oct 2011

Voting with Limited Information and Many Alternatives

Flavio Chierichetti * Jon Kleinberg †

October, 2011

Abstract

The traditional axiomatic approach to voting is motivated by the problem of reconciling differences in subjective preferences. In contrast, a dominant line of work in the theory of voting over the past 15 years has considered a different kind of scenario, also fundamental to voting, in which there is a genuinely "best" outcome that voters would agree on if they only had enough information. This type of scenario has its roots in the classical Condorcet Jury Theorem; it includes cases such as jurors in a criminal trial who all want to reach the correct verdict but disagree in their inferences from the available evidence, or a corporate board of directors who all want to improve the company's revenue, but who have different information that favors different options.

This style of voting leads to a natural set of questions: each voter has a *private signal* that provides probabilistic information about which option is best, and a central question is whether a simple plurality voting system, which tabulates votes for different options, can cause the group decision to arrive at the correct option. We show that plurality voting is powerful enough to achieve this: there is a way for voters to map their signals into votes for options in such a way that — with sufficiently many voters — the correct option receives the greatest number of votes with high probability. We show further, however, that any process for achieving this is inherently expensive in the number of voters it requires: succeeding in identifying the correct option with probability at least $1 - \eta$ requires $\Omega(n^3\epsilon^{-2}\log\eta^{-1})$ voters, where n is the number of options and ϵ is a distributional measure of the minimum difference between the options.

*Department of Computer Science, Cornell University, Ithaca NY 14853. Supported in part by NSF grant CCF-0910940.

†Department of Computer Science, Cornell University, Ithaca NY 14853. Supported in part by a John D. and Catherine T. MacArthur Foundation Fellowship, a Google Research Grant, a Yahoo! Research Alliance Grant, and NSF grants IIS-0705774, IIS-0910664, and CCF-0910940.

1 Introduction

Information-Based Voting. A dominant recent theme in the study of voting has been to trace differences in voters' preferences back to differences in the information they have about the world. This information-based approach has its roots in one of the earliest results in voting theory — the *Condorcet Jury Theorem*, which used the then-young theory of probability to model a situation in which a panel of jurors each wants to vote for the correct decision in a trial, but each juror may be wrong about what the correct decision is independently and with probability $p < \frac{1}{2}$ [17]. It is only very recently, however, that this approach has received deeper theoretical attention [2, 4, 6, 7, 16], leading to what is now a large and growing body of research.

The basic premise of the information-based approach to voting is that all voters want the best option for the group as a whole, but they disagree on what this best option is, based on the information they have. This models a wide range of situations where the differences among voters are not purely subjective, but instead based on uncertainty. For example, in most criminal trials the key question is genuinely whether the defendant committed the crime or not; all jurors want to reach the correct decision, but they disagree on which of the pieces of information presented at the trial are most salient. Similarly, all the members of a corporate board of directors may genuinely agree that the goal is to reach a decision that will most improve the company's future revenue, but they disagree on which course of action is most likely to achieve this. Even at the level of large populations, there can be cases where each voter wants a candidate whose election — for example — will lead to the strongest improvement in the economy, but there is disagreement among the voters about which candidate is most likely to achieve this.

This view of voters as information-processing agents trying to reach a correct decision has made it possible to develop models for a range of important phenomena in voting; these include the fact that voters realize they might be wrong, and the corollary that they can sometimes be convinced by evidence [2, 7], the corresponding role of deliberation in committee voting [10], and the fact that many voters may choose to abstain or not participate when they believe that others have more accurate information than they do [3, 8].

A Basic Model of Information-Based Voting. In this paper, we consider the following basic theoretical model that has received wide study [2, 7]. There is a decision to be made, involving selecting from among several possible *options* $A_1, \ldots, A_n$. One of these options is *correct*, and all voters want to select it. However, which option is correct is determined by a process that cannot be directly observed, and the voters have to use indirect signals to infer the correct option. Before casting a vote, each voter t receives a private *signal* equal to some value s_j, providing evidence for the identity of the correct option. (The full set of possible signals will be labeled $\{s_1, s_2, \ldots, s_C\}$.) We assume that certain kinds of signals are more plentiful when certain options are correct, and that voters know conditional probabilities of the form $\Pr[s_j \textit{ is received} \mid A_i \textit{ is correct}] = \rho_{ij}$. We further assume that no two options induce exactly the same set of conditional probabilities over signals. Based on the signal she receives, each voter casts a vote for one option, potentially using a randomized rule to map the signal to a vote. A *voting system* — a rule for mapping a collection of votes to a group decision — is then applied to these votes. We are interested in the probability that the group decision will be equal to the correct option A_i.

Much of the power of this model in economics and political science comes from the way in which it separates the *signals* received by the voters from the *options* they are voting on. This captures a basic property of voting in many real-life situation, including the ones described at the outset: the signals represent information and decision-making heuristics that the individual voters possess in their minds, while the options correspond to candidates or alternatives presented on a ballot. For many reasons, the institution of voting therefore does not (and generally cannot) consist of a simple sharing of everyone's signals. Instead, voters are only able to convey the information they possess in a more indirect fashion, by voting for one of the

given options. The crucial question is whether there is a (possibly randomized) algorithm each voter can apply to his or her signal to produce a vote, in such a way that the correct option is chosen.

For simplicity in the following discussion, we consider an equivalent formulation of this model (in the spirit of [1]), via an experiment involving urns and marbles. Suppose an experimenter has a collection of urns $A_1, \ldots, A_n$, and each urn contains marbles of colors $s_1, \ldots, s_C$. The fraction of marbles of color s_j in urn A_i is equal to ρ_{ij}; no two urns have exactly the same mixture of colors. Now, the experimenter announces to a set of test subjects that he is placing one of the urns $A_1, \ldots, A_n$ on a table. Each test subject draws and replaces a single marble from the urn on the table, without showing it to the other subjects, and then writes down a vote (on a secret ballot) for which urn she believes is on the table. The experimenter then applies a voting rule to these votes, producing a group decision, and awards the group a prize if this group decision is equal to the true urn that was on the table. It should be clear that this formulation is simply a rephrasing of the original model, with the urns representing the options and the marbles representing the signals.

Can Plurality Voting Produce the Correct Answer?. Much of the initial theoretical work on these issues focused on the case of $n = 2$ options — that is, voting when there are two alternatives, such as in a jury trial or a yes/no vote on a proposed rule [2, 4, 6, 7, 16]. But many settings involve more than two options, and in this case the following basic question has remained open. Suppose the votes will be aggregated simply using plurality voting, with the urn receiving the most votes chosen as the group decision. Is there a rule the voters can use for mapping signals (colors) to votes, such that for any instance of the problem with urns $A_1, \ldots, A_n$, a set of m voters will identify the correct urn with probability converging to 1 as m grows? And if so, how large a set of voters is needed to guarantee a success probability of $1 - \eta$, for a given $\eta > 0$?

Recent work has highlighted the challenge and general lack of understanding of this question with more than two options, raising it as an open problem and providing interesting results in highly structured special cases where the signal space is rich enough that each option has a disjoint set of one or more signals that uniquely favor it [12, 13]. For general sets of signals, the question has been open: if the signals are expressively weak compared to the full set of options, is there necessarily any strategy for mapping signals to votes that would lead to the correct outcome under a simple system like plurality voting?

Optimal Information-Based Voting: Main Results. Our first main result is that for any finite set of signals, and any finite set of options that induce distinct distributions over these signals, there is a strategy such that a sufficiently large set of voters can arrive at the correct option with high probability using plurality voting. In other words, each voter translates her signal into a vote in such a way that the option receiving the most votes is, with high probability, the correct one.

Second, we show that achieving this goal using plurality voting is very expensive: it requires a large number of voters. We give lower and upper bounds on the number of voters needed to achieve a high probability of correctness, parametrized by three quantities: the number of options, the number of signals, and a quantity measuring the minimum separation between the distributions over signals induced by any two options. The lower bound is the technically most involved of our results, and for two signals it is asymptotically tight in both the number of options and the separation parameter.

The technical core of our results is the case in which there are n options and 2 signals. Let ϵ be the minimum positive difference between the probability assigned to a fixed signal s_k by two different options i and j. With two signals, we show there is a strategy by which $O(n^3\epsilon^{-2}\log\eta^{-1})$ voters can arrive at the correct option using plurality voting with probability at least $1 - \eta$. The strategy is symmetric, in that all voters map signals to votes according to the same probabilistic rule. While the algorithm involves a carefully designed rule, it is based on a principle that is intuitively natural: the voters "hedge" against the possibility that their information points in the wrong direction, by sometimes choosing to vote for an option other than

the one supported by their signal. The bound achieved by our algorithm is tight: there are instances in which $\Omega(n^3\epsilon^{-2}\log\eta^{-1})$ voters are necessary to achieve such a guarantee; this lower bound applies even to asymmetric strategies in which different voters can use different rules.

Note that by the pigeonhole principle, the minimum difference ϵ is at most $1/(n-1)$, and hence ϵ^{-1} is a parameter that is at least as large as $n-1$. For example, the special case with urns $A_0, A_1, \ldots, A_n$, in which A_i contains i blue marbles and $n-i$ red marbles, has $\epsilon = 1/n$, and so for this problem the tight bound on the minimum number of voters needed is $\Theta(n^5 \log \eta^{-1})$.

A recurring theme in our results is this fifth-power dependence of the number of voters on n, in the case when ϵ^{-1} is close to n. As such, it is useful to provide some intuition at the outset for how this fifth-power dependence arises. Thus, the following description is deliberately informal, but gives a sense for where this functional form comes from. Let there be m voters, and for simplicity let us consider the special case from the previous paragraph, with urns $A_0, A_1, \ldots, A_n$, in which A_i contains i blue marbles and $n-i$ red marbles. Under the asymptotically optimal (randomized) algorithms we consider, the correct urn will receive a greater number of votes in expectation than any other urn; this is why, with enough voters, we will eventually be able to distinguish the correct urn using plurality voting. Now, we will find that the optimal algorithm has the following two properties. First, it spreads out the votes relatively uniformly across a set of $\Theta(n)$ urns, and so if there are m voters, each of the urns in this set receives $\Theta(m/n)$ votes in expectation. The second, subtler property is the crucial one: the optimal algorithm ensures that the correct urn receives the most votes in expectation using a delicate optimization under which the the expected number of votes received by the correct urn will exceed the expected number of votes received by the adjacent urns by a factor of only $1+\delta$, where $\delta = \Theta(n^{-2})$. As a result, to distinguish the correct urn with high probability, we need a number of samples that is sufficient to yield at least $\Theta(\delta^{-2}) = \Theta(n^4)$ votes for the correct urn. But since the correct urn receives only $\Theta(m/n)$ votes in expectation, this means that we need m to be $\Theta(n^5)$.

We observe that in our more general bound $O(n^3\epsilon^{-2}\log\eta^{-1})$, the form of the dependence on ϵ^{-1} is in fact necessary even if the voters could share their signals (rather than casting individual votes). Indeed, with $n = 2$ options that assign probabilities to signals differing by only ϵ, even a single observer would need to see $\Theta(\epsilon^{-2}\log\eta^{-1})$ signals in order to identify the correct option with probability at least $1-\eta$. Thus, with a constant number of options, plurality voting is allowing voters to aggregate their information with an efficiency that is within a constant factor of the efficiency achievable by a single person who could observe all signals directly.

For the case of $C > 2$ possible signals or colors, let ϵ denote the minimum ℓ_1 distance[1] of two distinct urns' probability distributions. We have an upper bound of $O\left((C\log C)^2 n^3 \epsilon^{-2} \log \frac{n}{\eta}\right)$ on the number of voters needed. Since the lower bound for the two-signal case applies with $C > 2$, it is tight in ϵ, and we lose only an exponentially small factor in n. Finding the correct dependence of the required number of voters on n and C is an interesting open question.

Under plurality voting, voters can only communicate the name of a single option in response to a signal. We also consider voting systems that allow voters to be much more expressive: *cumulative voting*, in which each vote consists of assigning a non-negative weight to each option (such that the weights sum to 1); and *Condorcet voting*, in which each vote consists of a ranking of all the options. For bichromatic urns, we show that cumulative voting requires only $O(\epsilon^{-2}\log\eta^{-1})$ voters in order to succeed with high probability; this is tight even compared to the baseline discussed above, when a single observer has access to all the signals. We show that a similar bound would hold for Condorcet voting, modulo an intriguing conjecture about distributions over permutations.

[1]We observe that, in the multicolor case, choosing the right parameter to define a notion of "distance" between urns is not as straightforward as in the bichromatic case. We chose ℓ_1 because it is the parameter that has been used in the literature to determine the minimum number of samples that allows an algorithm to distinguish between probability distributions.

Optimal Information-Based Voting: Main Techniques. The possibility result for identifying the correct option is based on a technique that implicitly draws a connection to the framework of *proper scoring rules* from statistics [11]. Proper scoring rules can be thought of as incentive systems for eliciting accurate probabilistic forecasts from expert predictors; the contexts in which they have been used in earlier work are quite different from ours, and to our knowledge there have not been previous linkages between proper scoring rules and information-based voting.

A construction based on proper scoring rules provides the first method for obtaining the correct option using plurality voting. However, we need to go beyond this construction in order to obtain a tight bound on the number of voters needed: in a sense to be made precise below, we can prove that any direct use of proper scoring rules in our setting requires at least $\Omega(n\epsilon^{-4})$ voters to achieve a high probability of success. This is at least $\Omega(n^3\epsilon^{-2})$ since $\epsilon \leq (n-1)^{-1}$, and more significantly, it has an asymptotically sub-optimal dependence on ϵ of $\Omega(\epsilon^{-4})$ when n is a constant, whereas our stronger approach achieves the optimal dependence of $\Theta(\epsilon^{-2}\log\eta^{-1})$ for constant n.

For the lower bound, we need to show that with $O(n^3\epsilon^{-2}\log\eta^{-1})$ voters, there is a probability η that plurality voting will choose the wrong option. For this, we identify a natural "close competitor" j of the correct option i, with a very similar signal distribution, and we consider a random variable that measures the extent to which the number of votes for the correct option i exceed the number for this competitor j. The (possibly asymmetric) strategies of the voters determine the variance of this random variable, and roughly speaking we follow a two-pronged argument in terms of this variance. If this variance is too low, then there is a high chance that voters would behave the same regardless of whether the option generating the signals was i or j, and hence that if they are correct about i with high probability, then they would have to be wrong with constant probability when j is the correct option. If the variance is above a certain low threshold, on the other hand, then we apply a carefully tuned "anti-concentration" inequality from [9, 14] showing that there is a constant probability that the number of votes for i will drop below the number for its competitor j.

Further Related Work. Finally, we mention two other recent lines of work that have also considered the problem faced by a set of agents trying to agree on a joint decision from a set of alternatives. Mossel, Sly, and Tamuz study a version of the problem in which there are two options, and each agent is given a probabilistic signal providing information about which option is correct [15]; in contrast to our approach and to the work on voting discussed above, they consider a model in which agents may communicate iteratively over multiple rounds. Caragiannis and Procaccia consider a setting based on agents that possess utilities over options; within this framework, they show that simple voting rules can approximately optimize the sum of agents' utilities for the option that is selected [5].

2 An Upper Bound with Two Signals

We begin by considering the case of two signals. Suppose we have a collection of n urns, labeled $p_1, \ldots, p_n$, the i-th of which having a p_i fraction of blue balls and a $1-p_i$ fraction of red balls, with $p_1 \leq p_2 \leq \cdots \leq p_n$. We let ϵ denote the smallest difference between two consecutive p_i's: $\epsilon = \min_{0\leq i\leq n-1}(p_{i+1} - p_i)$.

We assume that one urn is adversarially chosen as the correct one (we will also refer to this as the *unknown* urn). Then each player draws a ball from the urn and votes for the name of an urn based on the color they observe.

We describe the strategy that the players will use to randomly choose which vote to cast:

1. Let $b_k = \displaystyle\sum_{\ell=1}^{k-1} \frac{2-(p_{\ell+1}+p_\ell)}{p_{\ell+1}-p_\ell}$ and $r_k = \displaystyle\sum_{\ell=k}^{n-1} \frac{p_{\ell+1}+p_\ell}{p_{\ell+1}-p_\ell}$. Then define $R = \sum_{k=1}^{n} r_k$ and $B = \sum_{k=1}^{n} b_k$,

and set $M = \max(R, B)$.

2. The probability that a voter will vote for p_j if a red ball is drawn is $R_j = M^{-1} \cdot \left(r_j + \frac{M-R}{n}\right)$.

3. The probability that a voter will vote for p_j if a blue ball is drawn is $B_j = M^{-1} \cdot \left(b_j + \frac{M-B}{n}\right)$.

It is easy to check that the two given distribution are indeed probability distributions (their values are non-negative and they both sum up to one). Now, the probability that a player will vote for p_j given that the correct, adversarially chosen, distribution is p_i, is

$$\Pr_{\substack{\mathbf{X}\sim(p_i,1-p_i)\\ \mathbf{P}\sim f(\mathbf{X})}}[\mathbf{P} = p_j] = p_i \Pr_{\mathbf{P}\sim f(\text{blue})}[\mathbf{P} = p_j] + (1-p_i) \Pr_{\mathbf{P}\sim f(\text{red})}[\mathbf{P} = p_j] = p_i B_j + (1-p_i) R_j = E_i(j).$$

Now consider two urns, p_i and p_j. We compute the difference between the probabilities that a vote for urn p_i and a vote for urn p_j are cast, given that the correct urn is p_i:

$$\Delta_i(j) = E_i(i) - E_i(j) = p_i \left(B_i - B_j\right) + (1-p_i)\left(R_i - R_j\right).$$

We will lower-bound $\Delta_i(j)$ to bound the number of voters needed to let the voting scheme be successful with high probability. Suppose first that $j < i$; then

$$M \cdot \Delta_i(j) = p_i \cdot \sum_{\ell=j}^{i-1} \frac{2 - (p_{\ell+1} + p_\ell)}{p_{\ell+1} - p_\ell} - (1-p_i) \cdot \sum_{\ell=j}^{i-1} \frac{p_{\ell+1} + p_\ell}{p_{\ell+1} - p_\ell} = \sum_{\ell=j}^{i-1} \frac{2 \cdot p_i - (p_{\ell+1} + p_\ell)}{p_{\ell+1} - p_\ell},$$

observing that in each term of the sum we have $p_i \geq p_{\ell+1}$, since $\ell \leq i-1$. Therefore,

$$M \cdot \Delta_i(j) \geq \sum_{\ell=j}^{i-1} \frac{2p_{\ell+1} - (p_{\ell+1} + p_\ell)}{p_{\ell+1} - p_\ell} = \sum_{\ell=j}^{i-1} \frac{p_{\ell+1} - p_\ell}{p_{\ell+1} - p_\ell} = i - j.$$

If, instead, $i < j$ we have:

$$\begin{aligned} M \cdot \Delta_i(j) &= -p_i \cdot \sum_{\ell=i}^{j-1} \frac{2 - (p_{\ell+1} + p_\ell)}{p_{\ell+1} - p_\ell} + (1-p_i) \cdot \sum_{\ell=i}^{j-1} \frac{p_{\ell+1} + p_\ell}{p_{\ell+1} - p_\ell} \\ &= \sum_{\ell=i}^{j-1} \frac{(p_{\ell+1} + p_\ell) - 2p_i}{p_{\ell+1} - p_\ell} \geq \sum_{\ell=i}^{j-1} \frac{(p_{\ell+1} + p_\ell) - 2p_\ell}{p_{\ell+1} - p_\ell} = \sum_{\ell=i}^{j-1} \frac{p_{\ell+1} - p_\ell}{p_{\ell+1} - p_\ell} = j - i, \end{aligned}$$

where the inequality follows from $p_i \leq p_\ell$. Therefore for $j \neq i$, we have

$$\Delta_i(j) \geq \frac{|i-j|}{M}. \tag{1}$$

We now give an upper bound on the probability that the correct urn will be chosen by a voter. Note, somewhat counter-intuitively, that the probability of a correct vote is higher when this upper bound is smaller — this is because the $\Delta_i(j)$ are additive gaps, not multiplicative ones, and so by making the upper bound on the expected number of votes for the correct urn smaller, the gap $\Delta_i(j)$ becomes larger relative to the mean.

Recall that the correct urn is p_i. We upper-bound the probability that a vote will go to p_i:

$$\begin{aligned} E_i(i) &= p_i \cdot M^{-1} \cdot \left(b_i + \frac{M-B}{n}\right) + (1-p_i) \cdot M^{-1} \cdot \left(r_i + \frac{M-R}{n}\right) \\ &\leq p_i \cdot \left(M^{-1} \cdot b_i + \frac{1}{n}\right) + (1-p_i) \cdot \left(M^{-1} \cdot r_i + \frac{1}{n}\right). \end{aligned}$$

Observe that, by the definition of ϵ, we have that $b_i = \sum_{\ell=1}^{i-1} \frac{2-(p_{\ell+1}+p_\ell)}{p_{\ell+1}-p_\ell}$ satisfies $b_i \leq \frac{2i}{\epsilon}$, and furthermore that $r_i = \sum_{\ell=i}^{n-1} \frac{p_{\ell+1}+p_\ell}{p_{\ell+1}-p_\ell} \leq \frac{2(n-i)}{\epsilon}$. Thus,

$$E_i(i) \leq p_i \cdot \left(\frac{2i}{\epsilon M} + \frac{1}{n}\right) + (1-p_i) \cdot \left(\frac{2(n-i)}{\epsilon M} + \frac{1}{n}\right) \leq \frac{2n}{\epsilon M} + \frac{1}{n}.$$

We now give an upper bound on M. This will allow us to apply a Chernoff bound and finish the proof.

Recall that $M = \max(R, B)$; we will upper bound $R + B$ to get an upper bound on M:

$$R + B = \sum_{k=1}^{n}(r_k + b_k) \leq \sum_{k=1}^{n}(r_1 + b_n) = n \cdot (r_1 + b_n) = n \cdot \sum_{\ell=1}^{n-1} \frac{2}{p_{\ell+1} - p_\ell} \leq 2 \cdot \frac{n(n-1)}{\epsilon}.$$

It follows that

$$M \leq \frac{2n(n-1)}{\epsilon}. \tag{2}$$

Therefore, going back to the probability that an urn identical to the correct urn is voted for, we have

$$E_i(i) \leq \frac{2n}{\epsilon M} + \frac{1}{n} \leq \frac{2n}{\epsilon M} + \frac{1}{n} \cdot \frac{2n(n-1)}{M\epsilon} \leq \frac{4n}{\epsilon M}.$$

Furthermore, since $\Delta_i(j) > 0$ for each $j \neq i$, we have that the urn p_i is the most likely urn to be voted for, and therefore $E_i(i) \geq \frac{1}{n}$.

We are now ready to state the main theorem of the section. Its proof employs a careful application of the Chernoff bound, and the inequalities we have derived in this section.

Theorem 2.1. *Let urns $p_1, p_2, \ldots, p_n$ be given, with urn p_i having a p_i fraction of blue balls, and a $1 - p_i$ fraction of red balls. Let $0 \leq p_1 < p_2 < \cdots < p_n \leq 1$. Also, let ϵ be $\epsilon = \min_{1 \leq i \leq n-1}(p_{i+1} - p_i)$. Then, for Plurality Voting, $O\left(n^3\epsilon^{-2}\ln\eta^{-1}\right)$ voters are sufficient to guarantee a probability of at least $1 - \eta$ that the correct urn receives the most votes.*

Proof. Observe that $\epsilon \leq \frac{1}{n-1}$. Choose some $\eta \in (0, 1)$, and suppose the number of players is $m = \left\lceil 108 \cdot \frac{M(n-1)}{\epsilon} \cdot \ln\frac{4}{\eta} \right\rceil$ — we will show that m players will be enough to choose the correct option with probability is at least $1 - \eta$. Observe that, given our upper bound $M \leq \frac{2n(n-1)}{\epsilon}$, we have

$$m \leq \left\lceil 216 \cdot \frac{(n-1)^2 n}{\epsilon^2} \right\rceil.$$

Recall that we say that the players lose if an urn with a different distribution from the unknown urn wins the election. We will upper-bound the probability that the players lose, using the following form of the Chernoff bound:

Theorem 2.2 (Chernoff bound)**.** *Let $X_1, \ldots, X_m$ be independent $0/1$ random variables with expectation $E[X_i] = p_i$, for $i = 1, \ldots, n$. Let $\mu = \sum_i p_i$. Then, for each $\delta \geq 0$, it holds that*

$$\Pr\left[\sum_i X_i > (1+\delta) \cdot \mu\right] \leq \exp\left(-\frac{\min\left(\delta, \delta^2\right)}{3} \cdot \mu\right),$$

and,

$$\Pr\left[\sum_i X_i < (1-\delta)\cdot\mu\right] \le \exp\left(-\frac{\delta^2}{3}\cdot\mu\right).$$

We now show how to use Theorem 2.2, together with the bounds derived in Section 5.1, to prove Theorem 2.1. Let V_j be the number of votes to p_j in the random election, with unknown urn i. Then, $E[V_j] = E_i(j)\cdot m$. We have,

$$\begin{aligned}\Pr[\text{the players lose}] &= \Pr[p_i \text{ did not collect more votes than any other urn}]\\ &\le \Pr\left[V_i < E[V_i] - \frac{m}{3M}\right] + \sum_{\substack{j=0\\ j\neq i}}^{n} \Pr\left[V_j > E[V_i] - \frac{2m}{3M}\right]\end{aligned}$$

Since $\Delta_i(j) \ge \frac{|i-j|}{M}$ we have $E[V_i] \ge E[V_j] + \frac{|i-j|}{M}\cdot m$, and

$$\begin{aligned}\Pr[\text{the players lose}] &\le \Pr\left[V_i < E[V_i]\left(1 - \frac{m}{3ME[V_i]}\right)\right] + \sum_{\substack{j=0\\ j\neq i}}^{n} \Pr\left[V_j > E[V_j]\left(1 + \frac{|i-j|}{ME[V_j]}\cdot m - \frac{2m}{3ME[V_j]}\right)\right]\\ &\le \Pr\left[V_i < E[V_i]\left(1 - \frac{m}{3ME[V_i]}\right)\right] + \sum_{\substack{j=0\\ j\neq i}}^{n} \Pr\left[V_j > E[V_j]\left(1 + \frac{|i-j|\,m}{3ME[V_j]}\right)\right]\\ &\le \exp\left(-\frac{m^2}{27\,M^2\,E[V_i]}\right) + 2\sum_{k=1}^{n}\exp\left(-\min\left\{\frac{km}{9M}, \frac{k^2m^2}{27M^2E[V_j]}\right\}\right),\end{aligned}$$

by $E[V_j] \le E[V_i] \le m\cdot\frac{4(n-1)}{\epsilon M}$,

$$\begin{aligned}\Pr[\text{the players lose}] &\le \exp\left(-\frac{m^2\,\epsilon\,M}{108\,M^2\,m\,(n-1)}\right) + 2\sum_{k=1}^{n}\exp\left(-\min\left\{\frac{km}{9M}, \frac{k^2m^2\epsilon M}{108\,M^2\,m\,(n-1)}\right\}\right)\\ &\le \exp\left(-\frac{m\,\epsilon}{108\,M\,(n-1)}\right) + 2\sum_{k=1}^{n}\exp\left(-\min\left\{\frac{km}{9M}, \frac{k^2m\epsilon}{108\,M\,(n-1)}\right\}\right)\\ &\le \exp\left(-\frac{m\,\epsilon}{108\,M\,(n-1)}\right) + 2\sum_{k=1}^{n}\exp\left(-\frac{k^2m\epsilon}{108\,M\,(n-1)}\cdot\min\left\{\frac{12\,(n-1)}{k\epsilon}, 1\right\}\right)\\ &\le \exp\left(-\frac{m\,\epsilon}{108\,M\,(n-1)}\right) + 2\sum_{k=1}^{n}\exp\left(-k^2\ln\frac{4}{\eta}\right)\\ &\le \frac{\eta}{4} + 2\sum_{k=1}^{n}\left(\frac{\eta}{4}\right)^k \le \frac{\eta}{4} + 2\cdot\frac{\eta}{4-\eta} \le \frac{\eta}{4} + 2\cdot\frac{\eta}{3} < \eta.\end{aligned}$$

It follows that if $m = \Theta(n^3\epsilon^{-2}\log\eta^{-1})$, and voters apply the aforementioned voting scheme, the probability of winning is at least $1-\eta$. □

3 A Connection to Proper Scoring Rules

In this section we discuss the connection between our upper bound and the notion of a *proper scoring rule* [11]. We first show how to obtain a strategy for a set of voters in the two-signal case using proper

scoring rules.[2] We then show that basing a strategy on proper scoring rules cannot lead to an asymptotically tight result: any voting strategy based on the functions arising from the framework of proper scoring rules requires at least $\Omega\left(n\epsilon^{-4}\right)$ voters. This is weaker than the upper bound of $O(n^3\epsilon^{-2}\log\eta^{-1})$ that we obtained in Section 2 in two important respects. First, by the pigeon-hole principle, $\epsilon \leq \frac{1}{n-1}$, and therefore approach from the previous section is always at least as good as the approach based on proper scoring rules, and often much better. More significantly, when n is a constant, the approach via scoring rules gives a dependence on ϵ of $O(\epsilon^{-4})$, whereas our approach from Section 2 gives $O(\epsilon^{-2})$, which is optimal even if the group of voters could directly share all their signals. (In other words, even if there were just a single voter who received all the signals.)

For our purposes in this discussion, it is not necessary to introduce the full theory of proper scoring rules, but just to provide a self-contained consequence of that theory. The consequence is the following: it is possible to construct pairs of non-negative functions (f_0, f_1), each defined over the interval $[0, 1]$, with the property that for all $z \in [0, 1]$, the function

$$g_z(x) = zf_0(x) + (1 - z)f_1(x) \tag{3}$$

is uniquely maximized at $x = z$. We will further assume that f_0 and f_1 each have continuous second derivatives, which is true of the standard functions that arise from this theory. This defining property of f_0 and f_1 is all we will need.

From a pair of such functions, here is how we can define a strategy for each voter in the two-signal case. We have a set of $n + 1$ urns, where urn i has a probability p_i of producing a blue ball. We define $q_0 = \sum_i f_0(p_i)$ and $q_1 = \sum_i f_1(p_i)$, and let $q^* = \max(q_1, q_0)$. Now, when a voter draws a blue ball, they vote for urn i with probability proportional to $\dfrac{f_0(p_i)}{q^*} + \dfrac{q^* - q_0}{q^*(n+1)}$; if they draw a red ball, they vote for urn i with probability proportional to $\dfrac{f_1(p_i)}{q^*} + \dfrac{q^* - q_1}{q^*(n+1)}$. We call this the strategy *induced by f_0 and f_1*.

Suppose the true urn is t; then the number of votes for an urn j is a random variable $X_j = \sum_v X_{jv}$, where X_{jv} is the indicator variable that voter v votes for j. With k voters, we have

$$E\left[X_j\right] = \frac{k}{q^*}(p_t f_0(p_j) + (1 - p_t)f_1(p_j)) + \frac{|q_1 - q_0|\,k}{q^*(n+1)} = \frac{k}{q^*}g_{p_t}(p_j) + \frac{|q_1 - q_0|\,k}{q^*(n+1)}. \tag{4}$$

By the defining property of f_0 and f_1, we see that $E\left[X_j\right]$ is uniquely maximized at $j = t$. Hence for a sufficiently large set of voters, the number of votes received by urn t will exceed the number received by all other urns with high probability.

Thus, the strategy induced by any proper scoring rule will produce the true urn with high probability when there are enough voters. It can be viewed, in a sense, as a much simpler version of the construction in Section 2, and we now show that this simpler approach results in asymptotically larger number of voters.

Theorem 3.1. *Let f_0 and f_1 be any functions with continuous second derivatives that satisfy the defining property of proper scoring rules from Equation (3). Then in order for the strategy induced by f_0 and f_1 to identify the true urn with high probability, there must be $\Omega(n\epsilon^{-4})$ voters.*

Proof. We start with a basic claim about sums of Bernoulli trials. Let $X = \sum_{i^1}^{k} X_i$ be a sum of independent 0-1 random variables, where $E\left[X_i\right] = p_i \leq \frac{1}{2}$. The mean of X is $\mu = \sum_{i=1}^{k} p_i$. Then with constant probability, X will deviate by at least a constant multiple of $\sqrt{\operatorname{Var} X}$ from μ. More concretely, there are absolute constants $\alpha > 0$ and $\beta > 0$ so that with probability at least α, we have $X < \mu - \beta\sqrt{\operatorname{Var} X}$. Now, since

$$\operatorname{Var} X_i = p_i(1 - p_i) \geq p_i/2,$$

[2] We are grateful to Bobby Kleinberg for identifying this connection between voting strategies and proper scoring rules.

we have

$$\operatorname{Var} X = \sum_{i=1}^{k} \operatorname{Var} X_i \geq \sum_{i=1}^{k} p_i/2 \geq \mu/2.$$

Now, for a given $\delta > 0$, suppose we have $\mu < \beta^2/(2\delta^2)$. Then equivalently, $\delta < \beta/\sqrt{2\mu}$, so

$$\delta\mu < \beta\sqrt{\mu/2} \leq \beta\sqrt{\operatorname{Var} X}.$$

Hence with probability at least $\alpha > 0$, we have $X < (1-\delta)\mu$. It follows that in order to ensure $X \geq (1-\delta)\mu$ with probability going to 1, we must have $\mu \geq \Omega(1/\delta^2)$.

Now, recall that there are k voters, and consider the voting strategy induced by the functions f_0 and f_1. Since the first derivatives f_0' and f_1' are continuous functions defined over the compact set $[0,1]$, there is a constant c_1 such that $|f_0'(x)|, |f_1'(x)| \leq c_1$ for all $x \in [0,1]$. For the same reason, there is a constant c_2 such that $|f_0''(x)|, |f_1''(x)| \leq c_2$ for all $x \in [0,1]$. Using the bound on the first derivative, for any $\gamma > 0$ we can find an interval $[u,v] \subseteq [0,1]$ such that the following hold: (i) $d = \inf_{x \in [u,v]} \min(f_0(x), f_1(x)) > 0$, (ii) $v/u < 1+\gamma$, (iii) $(1-u)/(1-v) < 1+\gamma$,

$$\text{(iv)} \sup_{x,y \in [u,v]} \frac{f_0(y)}{f_0(x)} < 1+\gamma, \text{ and (v)} \sup_{x,y \in [u,v]} \frac{f_1(y)}{f_1(x)} < 1+\gamma.$$

It follows that if our probabilities $p_0, p_1, \ldots, p_n$ all lie in this interval $[u,v]$, then

$$E[X_j] \in \left[\frac{(1-\gamma_1)k}{n}, \frac{(1+\gamma_1)k}{n}\right]$$

for a constant γ_1 that goes to 0 with γ. Also, we have $q^* \geq dn$.

Now, for any $\epsilon > 0$, we choose $p_0 \leq p_1 \leq \cdots \leq p_n \in [u,v]$ such that $p_{j+1} - p_j = \epsilon$ for each j. Let t be the true urn, and let

$$h(x) = \frac{k}{q^*}(p_t f_0(x) + (1-p_t) f_1(x)) + \frac{|q_1 - q_0|\, k}{q^*(n+1)}.$$

Notice that $E[X_j] = h(p_j)$. Now, Taylor's Theorem implies that for some $w \in [p_t, p_{t+1}]$, we have

$$h(p_{t+1}) = h(p_t) + (p_{t+1} - p_t)h'(p_t) + \frac{1}{2}(p_{t+1} - p_t)^2 h''(w).$$

Since $h(x)$ has its global maximum at $x = p_t$, we have $h'(p_t) = 0$. Moreover, since $|f_0''(x)|, |f_1''(x)| \leq c_2$ for all $x \in [0,1]$, we have

$$h''(w) = \frac{k}{q^*}(p_t f_0''(x) + (1-p_t) f_1''(x)) \leq \frac{kc_2}{dn}.$$

Writing $p_{t+1} - p_t = \epsilon$, we have

$$h(p_{t+1}) \geq h(p_t) - \frac{kc_2}{2dn}\epsilon^2$$

Since $E[X_j] = h(p_j)$, for all j, this implies

$$E[X_{t+1}] \geq E[X_t] - \frac{kc_2}{2dn}\epsilon^2.$$

Since $E[X_t] \geq \frac{(1-\gamma_1)k}{n}$, this implies that $E[X_{t+1}] \geq (1-\delta)E[X_t]$, where $\delta = c_3\epsilon^2$ and $c_3 = \frac{c_2}{2(1-\gamma_1)d}$.

Now, using our initial fact about sums of Bernoulli trials, we must have $E[X_t] \geq \Omega(1/\delta^2)$ in order for X_t to have a high probability of exceeding $(1-\delta)E[X_t]$. Since $E[X_t] \leq \frac{(1+\gamma_1)k}{n}$, this requires

$$\frac{(1+\gamma_1)k}{n} \geq \frac{d_2}{c_3^2\epsilon^4}$$

for a constant $d_2 > 0$, and hence

$$k \geq \frac{d_2 n}{(1+\gamma_1)c_3^2\epsilon^4}.$$

□

4 A Tight Lower Bound for Two Signals

In this section we give a tight lower bound that confirms the optimality of the voting scheme for two signals presented in Section 2. We will start by introducing a class of instances. We will then prove a combinatorial lemma on how certain parameters of any (asymmetric) voting system for these instances have to behave, and we use the lemma to prove the lower bound.

We start by defining the lower bound class of instances $\mathcal{I}(n,\epsilon)$, for any $n \geq 2$ and $\epsilon \leq \frac{1}{n-1}$. The n urns in $\mathcal{I}(n,\epsilon)$ are such that $p_i = \frac{1-\epsilon(n-1)}{2} + (i-1)\epsilon$, for $i = 1,\ldots,n$. Then $0 \leq p_1 \leq p_2 \leq \cdots \leq p_n \leq 1$.

Each voter t is defined by two probability distributions $(R_{1,t}, R_{2,t}, \ldots, R_{n,t})$, $(B_{1,t}, B_{2,t}, \ldots, B_{n,t})$: if she draws a red (resp., blue) ball she will vote for urn i with probability $R_{i,t}$ (resp., $B_{i,t}$).

Given a voting scheme for m voters (that is, $2m$ probability vectors $(R_{i,t}), (B_{i,t})$), we define $B_i = m^{-1} \cdot \sum_{t=1}^{m} B_{i,t}$ and $R_i = m^{-1} \cdot \sum_{t=1}^{m} R_{i,t}$, for $i = 1,\ldots,m$. Thus the expected number of votes $E_i(j)$ to urn j, if i is the correct urn, will be equal to $m \cdot E_i(j) = m \cdot (p_i \cdot B_j + (1-p_i) \cdot R_j)$. We also define $\Delta_i(j) = E_i(i) - E_i(j)$ to be the expected difference between the number of votes to i and j, if i is the correct urn, averaged over the m voters.

We say that a voting scheme is *proper* if $\Delta_i(j) \geq 0$, for each i, j. The challenge in proving the lower bound lies in the fact that proper voting schemes can succeed in identifying the correct urn for what seem to be a variety of different reasons, and so we need to find a common property they have which implies that the correct urn only "narrowly" wins the election over other urns with very similar distributions. This is the content of the following lemma.

Lemma 4.1. *Let n and ϵ be $n \geq 10$ and $\epsilon \leq \frac{1}{n-1}$. Then all proper voting schemes for $\mathcal{I}(n,\epsilon)$ satisfy:*

(a) $B_1 \leq B_2 \leq \cdots \leq B_n \leq \frac{9}{n}$ *and* $\frac{9}{n} \geq R_1 \geq R_2 \geq \cdots \geq R_n$*;*

(b) $E_i(i) \leq \frac{9}{n}$*, for* $i = 1,\ldots,n$*.*

(c) There exists a set $S \subseteq [n]$ and $\iota \in \{-1,+1\}$, with $|S| \geq \frac{n}{4} - 3\ln n - 14$, such that for each $i \in S$, we have

$$\max\left(|R_i - R_{i+\iota}|, |B_i - B_{i+\iota}|\right) < e^{\frac{7}{2}} \cdot \frac{\sqrt{|R_i - B_i|}}{n^{\frac{3}{2}}},$$

and

$$\Delta_i(i+\iota), \Delta_{i+\iota}(i) \leq 2e^{\frac{7}{2}} \cdot \epsilon \cdot \frac{\sqrt{|R_i - B_i|}}{n^{\frac{3}{2}}}.$$

The crux of the lemma is to show that for many pairs of urns $i, i+\iota$, the election will be very "close": if i is the correct urn, it does not win the election by a large margin over $i+\iota$ in expectation (and vice versa). The lemma shows further that, averaged over the voters, the difference between the probability of voting for i given a red (resp., blue) ball and the probability of voting for $i+\iota$ given the same color is small.

This upper bound is crucial for the proof of the lower-bound theorem, stated next: we will show that — even if we only cared about urns $i, i+\iota$ — the variance of a voter's choice can be lower-bounded by $\Omega(|R_{i,t} - B_{i,t}|)$. This, assuming that the total variance is at least some constant, will allow us to apply an anti-concentration inequality to show that the expected margin $\Delta_i(i+\iota)$ of urn i over urn $i+\iota$ will be surpassed by $\Theta\left(\ln \eta^{-1}\right)$ standard deviations of the number of votes to urn i and $i+\iota$. It will follow that with probability $\Omega(\eta)$ the election will be won by the wrong urn. Again, this argument requires that the variance be at least some sufficiently large constant; if the variance is actually smaller than this constant, we will use a different argument showing that the voting system is sufficiently "inflexible" that if urn i wins when it is correct, the same pattern of votes is likely to also arise — favoring i — when $i+\iota$ is actually correct.

Proof. We will first show that in a proper voting scheme, for each $i < j$ it holds that $B_i \le B_j$ and $R_i \ge R_j$. This implies $B_0 \le B_1 \le \cdots \le B_{n-1}$ and $R_0 \ge R_1 \ge \cdots \ge R_{n-1}$. By contradiction,

- if $B_i < B_j$ and $R_i < R_j$, then $E_k(j) > E_k(i)$ for each k: in particular for $k = i$, which would give $\Delta_i(j) < 0$, contradicting the properness of the voting scheme;
- the same argument gives a contradiction if $B_i > B_j$ and $R_i > R_j$ (choosing $k = j$);
- finally, assume $B_i > B_j$, $R_i < R_j$, and that $E_i(i) > E_i(j)$, $E_j(i) < E_j(j)$; then,

$$\begin{gathered} E_i(i) - E_j(i) > E_i(j) - E_j(j) \\ (p_i - p_j)B_i + (p_j - p_i)R_i > (p_i - p_j)B_j + (p_j - p_i)R_j \\ (p_i - p_j)(B_i - B_j) > (p_j - p_i)(R_j - R_i), \end{gathered}$$

  then, by $p_i < p_j$, we have

$$(R_j - R_i) + (B_i - B_j) < 0,$$

  which is impossible since the left-hand side is positive by $B_i > B_j$ and $R_j > R_i$.

It follows that $B_i \le B_j$ and $R_i \ge R_j$. We now show that $B_n \le \frac{9}{n}$ (resp., $R_1 \le \frac{9}{n}$). Since $\{B_i\}_{i=1}^n$ and $\{R_i\}_{i=1}^n$ are probability distributions, one has that $\left|\left\{i \mid B_i + R_i \le \frac{4}{n}\right\}\right| \ge \frac{n}{2}$, for otherwise

$$2 = \sum_{i=1}^{n} B_i + \sum_{i=1}^{n} R_i \ge \sum_{\substack{i=1 \\ B_i + R_i > \frac{4}{n}}}^{n} E_i(i) > \frac{n}{2} \cdot \frac{4}{n} \ge 2.$$

Since $p_{\lfloor \frac{n+1}{2} \rfloor} \ge \frac{1-\epsilon}{2}$ (resp., $p_{\lceil \frac{n+1}{2} \rceil} \le \frac{1+\epsilon}{2}$), it follows that there exists some i such that $B_i + R_i \le \frac{4}{n}$ and $p_i \ge \frac{1-\epsilon}{2}$ (resp., $p_i \le \frac{1+\epsilon}{2}$). Observe that $E_i(i) \le \frac{4}{n}$. Now, by contradiction, let $B_n > \frac{9}{n}$ ($R_1 > \frac{9}{n}$); by $n \ge 10$ we have $\epsilon \le \frac{1}{n-1} \le \frac{1}{9}$; therefore $p_i \ge \frac{4}{9}$ ($p_i \le \frac{5}{9}$) one has $E_i(n) \ge p_i \cdot B_n > \frac{4}{n}$ ($E_i(1) \ge p_i \cdot R_1 > \frac{4}{n}$). It follows that $\Delta_i(n)$ ($\Delta_i(1)$) is negative, contradicting properness.

We define $\delta_i = |R_i - B_i|$. Then $\delta_n \le B_n \le \frac{9}{n}$ and $\delta_1 \le R_1 \le \frac{9}{n}$. Let k_r be the largest integer such that $R_{k_r+1} \ge B_{k_r+1}$, and k_b be the largest integer such that $B_{n-k_b} \ge R_{n-k_b}$. Observe that (a) if $i \le k_r$ then $\delta_i \ge \delta_{i+1}$, (b) if $i \ge n - k_b + 1$ then $\delta_i \ge \delta_{i-1}$, and (c) $k_r + k_b \ge n - 2$. By (c) at least one of k_r and k_b

has to be at least $\frac{n}{2}-1$. We let S_R and S_B be

$$S_R = \left\{ i \mid R_{i+1} \geq B_{i+1} \wedge \max\left(R_i - R_{i+1}, B_{i+1} - B_i\right) < e^{7/2} \cdot \frac{\sqrt{R_i - B_i}}{n^{3/2}} \right\},$$
$$S_B = \left\{ i \mid B_{i-1} \geq R_{i-1} \wedge \max\left(R_{i-1} - R_i, B_i - B_{i-1}\right) < e^{7/2} \cdot \frac{\sqrt{B_i - R_i}}{n^{3/2}} \right\}.$$

Then $S_R \subseteq \{1, 2, \ldots, k_r\}$ and $S_B \subseteq \{n - k_b + 1, n - k_b + 2, \ldots, n-1, n\}$. We consider two cases:

- suppose $k_r \geq \frac{n}{2} - 1$. We relabel the element in $\overline{S}_R = [k_r] - S_R$, using $r = \left|\overline{S}_R\right|$:

  $$\overline{S}_R = \{i_1, i_2, \ldots, i_r\},$$

  with $i_1 < i_2 < \cdots < i_r$. We have $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_{k_r} \geq \delta_{k_r+1}$. Then, for $1 \leq t \leq r-1$,

  $$\delta_{i_{t+1}} \leq \delta_{i_t+1} \leq \delta_{i_t} - e^{7/2} \cdot \frac{\sqrt{R_{i_t} - B_{i_t}}}{n^{3/2}} = \delta_{i_t} - e^{7/2} \cdot \frac{\sqrt{\delta_{i_t}}}{n^{3/2}} = \delta_{i_t} \cdot \left(1 - \sqrt{\frac{e^7}{n^3 \cdot \delta_{i_t}}}\right), \quad (5)$$

  and $\delta_{i_1} \leq \frac{9}{n} \leq \frac{e^3}{n}$; we define $\alpha_k = e^{3-k}$ so that $\delta_{i_1} \leq \alpha_0 \cdot n^{-1}$. Let $\ell_0 = 1$ and $\ell_k = \left\lceil n \cdot e^{-(k+3)/2} \right\rceil$, for $k \geq 1$. We also let $L(k) = \sum_{j=0}^{k} \ell_j$. We will show by induction on k that $\delta_{i_{L(k)}} \leq \alpha_k \cdot n^{-1} = e^{3-k} \cdot n^{-1}$. The case $k = 0$ has already been verified. We assume $k \geq 1$. Then,

  $$\begin{aligned}
  \delta_{i_{L(k)}} &\leq \delta_{i_{L(k)-1}} \cdot \left(1 - \sqrt{\frac{e^7}{n^3 \cdot \delta_{i_{L(k)-1}}}}\right) \leq \delta_{i_{L(k-1)}} \cdot \left(1 - \sqrt{\frac{e^7}{n^3 \cdot \delta_{i_{L(k-1)}}}}\right)^{\ell_k} \\
  &\leq \delta_{i_{L(k)-1}} \cdot \left(1 - \sqrt{\frac{e^{k+3}}{n^2}}\right)^{\ell_k} = \delta_{i_{L(k-1)}} \cdot \left(1 - \frac{e^{\frac{k+3}{2}}}{n}\right)^{\ell_k} \\
  &= \delta_{i_{L(k-1)}} \cdot \left(1 - \frac{e^{\frac{k+3}{2}}}{n}\right)^{\left\lceil \frac{n}{e^{(k+3)/2}} \right\rceil} \leq \delta_{i_{L(k-1)}} \cdot e^{-1} \leq e^{3-k} \cdot n^{-1}.
  \end{aligned}$$

  Now, if $k \geq 11 + 3\ln n$, we have $\delta_{i_{L(k)}} \leq n^3 \cdot e^{-8}$; by (5), we would then get $\delta_{i_{L(k)+1}} \leq \delta_{i_{L(k)}} \cdot (1-e) < 0$ — since $\delta_{i_r} \geq 0$, by $i_r \leq k_r$, this implies that $r = \left|\overline{S}_R\right| < L\left(\lceil 11 + 3\ln n \rceil\right)$. We now upper bound $L(k)$ to get an upper bound on $r = \left|\overline{S}_R\right|$:

  $$\begin{aligned}
  L(k) &= \sum_{j=0}^{k} \ell_j = 1 + \sum_{j=1}^{k} \left\lceil n \cdot e^{-\frac{k+3}{2}} \right\rceil \\
  &= k + 1 + n \cdot \sum_{j=1}^{k} e^{-\frac{k+3}{2}} \leq k + 1 + n \cdot e^{-5/2} \cdot \sum_{j=0}^{\infty} e^{-k/2} \\
  &= k + 1 + n \cdot e^{-5/2} \cdot \frac{1}{1 - e^{-1/2}} = k + 1 + \frac{n}{e^{5/2} - e^2} \leq \frac{n}{4} + k + 1.
  \end{aligned}$$

  It follows that $r = \left|\overline{S}_R\right| < L\left(\lceil 11 + 3\ln n \rceil\right) \leq \frac{n}{4} + \lceil 11 + 3\ln n \rceil + 1 \leq \frac{n}{4} + 3\ln n + 13$, and therefore

  $$|S_R| \geq k_r - \frac{n}{4} - 3\ln n - 13 \geq \frac{n}{4} - 3\ln n - 14.$$

- Otherwise, $k_r < \frac{n}{2} - 1$ and therefore $k_b > \frac{n}{2} - 1$. A proof similar to the previous case gives

$$|S_B| \geq k_b - \frac{n}{4} - 3\ln n - 13 \geq \frac{n}{4} - 3\ln n - 14.$$

Therefore, at least one of S_R and S_R has cardinality at least $\frac{n}{4} - 3\ln n - 14$. If S_R is the largest one we pick $\iota = 1$ and $S = S_R$. Otherwise, we pick $\iota = -1$ and $S = S_B$. Observe that the choice satisfies the first requirement of point (c) in the statement.

We now prove the second requirement of point (c). Let i be an element of S, $\beta = B_i - B_{i+\iota}$, and $\rho = R_i - R_{i+\iota}$. Then, $|\beta| = -\iota\beta$ and $|\rho| = \iota\rho$. Also,

$$|\beta|, |\rho| \leq e^{\frac{7}{2}} \cdot \frac{\sqrt{|R_i - B_i|}}{n^{\frac{3}{2}}}.$$

Recall that $\Delta_i(i+\iota) = E_i(i) - E_i(i+\iota) = \beta p_i + \rho(1-p_i)$ and $\Delta_{i+\iota}(i) = E_{i+\iota}(i+\iota) - E_{i+1}(i) = -\beta p_{i+\iota} - \rho(1-p_{i+\iota})$. Suppose that at least one of $\Delta_i(i+\iota)$ and $\Delta_{i+\iota}(i)$ is larger than $2e^{\frac{7}{2}}\epsilon\frac{\sqrt{|R_i-B_i|}}{n^{\frac{3}{2}}}$. By the properness of the voting system, we would have:

$$\Delta_i(i+\iota) + \Delta_{i+\iota}(i) > 2e^{\frac{7}{2}} \cdot \epsilon \cdot \frac{\sqrt{|R_i - B_i|}}{n^{\frac{3}{2}}}$$
$$\beta(p_i - p_{i+\iota}) + \rho(p_{i+\iota} - p_i) > 2e^{\frac{7}{2}} \cdot \epsilon \cdot \frac{\sqrt{|R_i - B_i|}}{n^{\frac{3}{2}}},$$

by the definition of the instance, we have that $p_{i+\iota} - p_i = \iota \cdot \epsilon$, therefore we would have

$$|\beta| + |\rho| > 2e^{\frac{7}{2}} \cdot \frac{\sqrt{|R_i - B_i|}}{n^{\frac{3}{2}}}$$

which would imply that at least one of $|\beta|$ and $|\rho|$ is larger than $e^{\frac{7}{2}} \cdot \frac{\sqrt{|R_i-B_i|}}{n^{\frac{3}{2}}}$, a contradiction. It follows that both $\Delta_i(i+\iota)$ and $\Delta_{i+\iota}(i)$ are less than or equal

$$\Delta_i(i+\iota), \Delta_{i+\iota}(i) \leq 2e^{\frac{7}{2}} \cdot \epsilon \cdot \frac{\sqrt{|R_i - B_i|}}{n^{\frac{3}{2}}}.$$

□

Theorem 4.2. *There exists a positive constant H such that for any $\eta < H$, one has that any voting scheme for $\mathcal{I}(n,\epsilon)$, with $n \geq 120$ and $\epsilon \leq \frac{1}{11(n-1)}$ using at most $O\left(\frac{n^3}{\epsilon^2}\log\frac{1}{\eta}\right)$ voters will fail to win the election with probability $\Omega(\eta)$.*

Proof. Take any asymmetric voting scheme for $\mathcal{I}(n,\epsilon)$ with m voters — that is, a sequence of m vectors $(R_{1,t},\ldots,R_{n,t})$ and $(B_{1,t},\ldots,B_{n,t})$, for $1 \leq t \leq m$, such that the probability that the tth voter votes for the ith urn if she draws a blue (resp., red) ball is $B_{i,t}$ (resp., $R_{i,t}$). Let $B_i = m^{-1} \cdot \sum_{t=1}^{m} B_{i,t}$ and $R_i = m^{-1} \cdot \sum_{t=1}^{m} R_{i,t}$.

If the voting scheme is improper, then by definition there exists i, j such that $\Delta_i(j) < 0$. Otherwise, by $n \geq 120$, one has $\frac{n}{4} - 3\ln n - 14 \geq 1$, and by Lemma 4.1, there will exist two urns i and $j \in \{i-1, i+1\}$ such that $\Delta_i(j) \leq 2e^{\frac{7}{2}}\epsilon\frac{\sqrt{|R_i-B_i|}}{n^{\frac{3}{2}}}$.

Given i, j, we define the head-to-head (i,j)-voting process as follows; for each voter t, the random variable $X_t = X_t(i,j)$ will be defined as

$$X_t = \begin{cases} 1 & \text{if voter } t \text{ votes for urn } j, \text{ given that the unknown urn is } i, \\ 1/2 & \text{if voter } t \text{ does not vote for urns } i \text{ or } j, \text{ given that the unknown urn is } i, \\ 0 & \text{if voter } t \text{ votes for the unknown urn } i. \end{cases}$$

Observe that $X = \sum_{t=1}^{m} X_t \geq \frac{m}{2}$ iff the number of votes to urn j is not smaller than the number of votes to the right urn i. In this case, the voters will lose the election. Furthermore,

$$E[X] = \frac{m}{2} - \frac{m}{2} \cdot \Delta_i(j).$$

Since X is the sum of independent random variables, we have that $\mathrm{Var}[X] = \sum_{t=1}^{m} \mathrm{Var}[X_t]$; by $\epsilon \leq \frac{1}{11}(n-1)$, we have that $\frac{5}{11} \leq p_1 \leq p_2 \leq \cdots \leq p_n \leq \frac{6}{11}$. We will use that $p_i, 1-p_i \geq \frac{1}{4}$ for each i, to lower-bound the variance of X_t:

$$\begin{aligned} \mathrm{Var}[X_t] = {} & (p_i \cdot B_{i,t} + (1-p_i) \cdot R_{i,t}) \cdot (0 - E[X_t])^2 + (p_i \cdot B_{j,t} + (1-p_i) \cdot R_{j,t}) \cdot (1 - E[X_t])^2 \\ & + (p_i \cdot (1 - B_{i,t} - B_{j,t}) + (1-p_i) \cdot (1 - R_{i,t} - R_{j,t})) \cdot \left(\frac{1}{2} - E[X_t]\right)^2. \end{aligned}$$

We consider two cases:

- if $E[X_t] \geq \frac{1}{4}$, then

$$\mathrm{Var}[X_t] \geq (p_i B_{i,t} + (1-p_i) R_{i,t}) \cdot (0 - E[X_t])^2 \geq \frac{R_{i,t} + B_{i,t}}{4} \cdot \frac{1}{16} \geq \frac{|R_{i,t} - B_{i,t}|}{64}.$$

- if $E[X_t] < \frac{1}{4}$, we have

$$\begin{aligned} \mathrm{Var}[X_t] & \geq (p_i B_{j,t} + (1-p_i) R_{j,t}) (1 - E[X_t])^2 \\ & \quad + (p_i(1 - B_{i,t} - B_{j,t}) + (1-p_i)(1 - R_{i,t} - R_{j,t})) \left(\frac{1}{2} - E[X_t]\right)^2 \\ & \geq \frac{p_i B_{j,t} + (1-p_i) R_{j,t}}{16} + \frac{p_i(1 - B_{i,t} - B_{j,t}) + (1-p_i)(1 - R_{i,t} - R_{j,t})}{16} \\ & = \frac{p_i(1 - B_{i,t}) + (1-p_i)(1 - R_{i,t})}{16}. \end{aligned}$$

  The latter is equal to both $\frac{1}{16} p_i (R_{i,t} - B_{i,t}) + (1 - R_{i,t})$ and $\frac{1}{16}(1-p_i)(B_{i,t} - R_{i,t}) + (1 - B_{i,t})$; we can therefore get a lower bound of

$$\mathrm{Var}[X_t] \geq \frac{1}{16} \cdot \min(p_i, 1-p_i) \cdot |R_{i,t} - B_{i,t}| \geq \frac{|R_{i,t} - B_{i,t}|}{64}$$

It follows that $\mathrm{Var}[X] = \sum_{t=1}^{m} \mathrm{Var}[X_t] \geq \frac{1}{64} \cdot \sum_{t=1}^{m} |R_{i,t} - B_{i,t}|$.

Recall that $m \cdot R_i = \sum_{t=1}^{m} R_{i,t}$ and $m \cdot B_i = \sum_{t=1}^{m} B_{i,t}$. Suppose $R_i \geq B_i$; then

$$m \cdot |R_i - B_i| = m \cdot (R_i - B_i) = m \cdot \sum_{t=1}^{m} (R_{i,t} - B_{i,t}) \leq m \cdot \sum_{t=1}^{m} |R_{i,t} - B_{i,t}| \leq 64 \cdot \mathrm{Var}[X].$$

If, on the other hand, $B_i > R_i$, we have

$$m \cdot |R_i - B_i| = m \cdot (B_i - R_i) = m \cdot \sum_{t=1}^{m} (B_{i,t} - R_{i,t}) \leq m \cdot \sum_{t=1}^{m} |B_{i,t} - R_{i,t}| \leq 64 \cdot \mathrm{Var}[X].$$

Therefore, in any case, we have $\mathrm{Var}[X] \geq \frac{m|R_i - B_i|}{64}$.

We now give a different lower bound on $\mathrm{Var}[X]$, that we will use to deal with the case of very small variance $\mathrm{Var}[X]$. Let $p_{1,t}, p_{1/2,t}, p_{0,t}$ be, respectively, the probabilities that $X_t = 1, X_t = \frac{1}{2}$ and $X_t = 0$. Then, $E[X_t] = p_{1,t} + \frac{1}{2} \cdot p_{1/2,t}$, and

$$\mathrm{Var}[X_t] = p_{1,t} \cdot (E[X_t] - 1)^2 + p_{1/2,t} \cdot \left(E[X_t] - \frac{1}{2} \right)^2 + p_{0,t} \cdot (E[X_t])^2$$

We consider three cases:

- if $p_{1,t} = \max\left(p_{1,t}, p_{1/2,t}, p_{0,t}\right) \geq \frac{1}{3}$ then if $E[X_t] \leq \frac{3}{4}$ we have

  $$\mathrm{Var}[X_t] \geq p_{1,t} \cdot (E[X_t] - 1)^2 \geq \frac{1}{3} \cdot \frac{1}{4^2} = \frac{1}{48} \geq \frac{1 - p_{1,t}}{48}.$$

  If instead $E[X_t] > \frac{3}{4}$, then

  $$\mathrm{Var}[X_t] \geq p_{1/2,t} \left(E[X_t] - \frac{1}{2} \right)^2 + p_{0,t} \left(E[X_t] \right)^2 > \frac{p_{1/2,t}}{4^2} + \frac{p_{0,t}}{4^2} \geq \frac{1 - p_{1,t}}{16}.$$

- If $p_{0,t} = \max\left(p_{1,t}, p_{1/2,t}, p_{0,t}\right) \geq \frac{1}{3}$ then we employ a similar approach. If $E[X_t] \geq \frac{1}{4}$ we have

  $$\mathrm{Var}[X_t] \geq p_{0,t} \cdot (E[X_t])^2 \geq \frac{1}{3} \cdot \frac{1}{4^2} = \frac{1}{48} \geq \frac{1 - p_{0,t}}{48}.$$

  If $E[X_t] < \frac{1}{4}$, then

  $$\mathrm{Var}[X_t] \geq p_{1/2,t} \left(E[X_t] - \frac{1}{2} \right)^2 + p_{1,t} \left(E[X_t] - 1 \right)^2 \geq \frac{p_{1/2,t}}{4^2} + \frac{p_{1,t}}{4^2} \geq \frac{1 - p_{0,t}}{16}.$$

- If $p_{1/2,t} = \max\left(p_{1,t}, p_{1/2,t}, p_{0,t}\right) \geq \frac{1}{3}$, then $\frac{1}{6} \leq E[X_t] \leq \frac{5}{6}$, and

  $$\mathrm{Var}[X_t] \geq p_{1,t} \left(E[X_t] - 1 \right)^2 + p_{0,t} \left(E[X_t] \right)^2 \geq \frac{p_{1,t}}{6^2} + \frac{p_{0,t}}{6^2} \geq \frac{1 - p_{1/2,t}}{36}.$$

In each of the three cases, we had $\mathrm{Var}[X_t] \geq \frac{1 - \max\left(p_{1,t}, p_{1/2,t}, p_{0,t}\right)}{48}$, and therefore

$$\mathrm{Var}[X] = \sum_{t=1}^{m} \mathrm{Var}[X_t] \geq \frac{1}{48} \cdot \sum_{t=1}^{m} \left(1 - \max\left(p_{1,t}, p_{1/2,t}, p_{0,t}\right) \right).$$

Let us now assume that $\mathrm{Var}[X] \leq \frac{1}{72} \cdot \log_5 \frac{1}{\eta}$. We will deal with the case $\mathrm{Var}[X] > \frac{1}{72} \cdot \log_5 \frac{1}{\eta}$ later. The previous inequality then implies

$$\sum_{t=1}^{m} \left(1 - \max\left(p_{1,t}, p_{1/2,t}, p_{0,t}\right) \right) \leq \frac{2}{3} \log_5 \frac{1}{\eta}.$$

Recall that $X = X(i,j) \geq \frac{m}{2}$ iff the unknown urn i gets at most as many votes as j (and therefore the election is lost). In the following we will also consider $X' = X(j,i)$; we have that $X' \leq \frac{m}{2}$ iff urn i gets at least as many votes as the unknown urn j (this also implies that the election is lost).

Observe that, since $\frac{5}{11} \leq p_1 \leq p_2 \leq \cdots \leq p_n \leq \frac{6}{11}$, no matter what the unknown urn is, the probability that any specific voter votes for any specific urn changes by a constant factor (between $\frac{5}{6}$ and $\frac{6}{5}$) if one changes the unknown urn.

We now show that, given that $\mathrm{Var}[x] \leq \frac{1}{72} \cdot \log_5 \eta^{-1}$, then with probability at least $\eta/9$ each voter will vote according to its maximum probability choice: that is

$$\Pr[\forall t, X_t \text{ equals the value } x_t \text{ that maximizes } \Pr[X_t = x_t]] \geq \frac{\eta}{9}.$$

If these choices let an urn different from i win the election, we have proven the theorem. Otherwise, we show that — if we exchange the unknown urn with any other urn k — then still with probability at least $\eta/25$ each voter t will vote for the same urn x_t; implying either a tie at the top, or that i (which would then not be the correct urn anymore) would will the election.

We let s_t denote the sum of the two minimum probabilities in $\{p_{1,t}, p_{1/2,t}, p_{0,t}\}$; that is $s_t = 1 - \max\left(p_{1,t}, p_{1/2,t}, p_{0,t}\right)$. Observe that $s_t \leq 1 - \frac{1}{3} = \frac{2}{3}$ for each t. If we define $s = \sum_{t=1}^{m} s_t$, we also have $s \leq \frac{2}{3}\log_5 \eta^{-1}$.

We have,

$$\Pr\left[\forall t, X_t \text{ equals the value } x_t \text{ that maximizes } \Pr[X_t = x_t]\right] = \prod_{t=1}^{m}(1 - s_t).$$

We now lower-bound the product, using the following greedy algorithm: take one of the largest $s_{t'} < \frac{2}{3}$, and one of the smallest $s_t > 0$, with $s_t \neq s_{t'}$. Then move $x = \min\left(\frac{2}{3} - s_{t'}, s_t\right) > 0$ mass from s_t to $s_{t'}$. Observe that the sum s of the s_t's remains constant throughout the process; furthermore the product $\prod_{m=1}^{t} s_t$ decreases: indeed, consider the product of $s_t \cdot s_{t'}$ before and after the change — we can disregard the rest since it remains constant. Let $s_t, s_{t'}$ be the two values before the change, and $s_t - x, s_{t'} + x$ be the two values after the change. That product used to be $s_t \cdot s_{t'}$, and becomes $s_t \cdot s_{t'} - x(s_{t'} - s_t) - x^2$ — the latter is smaller than $s_t \cdot s_{t'}$ since $x > 0$ and $s_{t'} > s_t$. Note also that at each step one of the s_t's stops being considered (either because it becomes equal to $\frac{2}{3}$ or equal to 0) — therefore the algorithm terminates. At termination there will exist at most one s_t with value different from $\frac{2}{3}$ and 0. Furthermore, recalling that $s = \sum_{t=1}^{m} s_t$, we conclude that then there will exist exactly $\left\lceil \frac{s}{2/3} \right\rceil$ different s_t's with value $2/3$, one with value $0 \leq s - \left\lceil \frac{s}{2/3} \right\rceil \cdot \frac{2}{3} < \frac{2}{3}$, and all the others having null value.

Given that $s \leq \frac{2}{3}\log_5 \eta^{-1}$, we can then minimize the former probability with

$$\begin{aligned}
\Pr[\forall t, X_t \text{ equals the value } x_t \text{ that maximizes } \Pr[X_t = x_t]] &= \prod_{t=1}^{m}(1 - s_t) \\
&\geq 3^{-\left\lceil \frac{s}{2/3} \right\rceil - 1} \\
&\geq 3^{-\left\lceil \log_5 \eta^{-1} \right\rceil - 1} \\
&\geq \frac{1}{9} \cdot 3^{-\log_5 \eta^{-1}} \\
&\geq \frac{1}{9}\eta^{\frac{1}{\log_3 5}} \geq \frac{\eta}{9}.
\end{aligned}$$

If these sequence of votes guarantees that the unknown urn i loses the election, we are done. Otherwise, we exchange the roles of urns i and j.

Recall that $\frac{5}{6} \leq \frac{p_i}{p_j}, \frac{1-p_i}{1-p_j} \leq \frac{6}{5}$ — and therefore, for each t, $s'_t \leq \frac{6}{5}s_t \leq \frac{4}{5}$. Indeed, let $\{a, b, c\} = \{1, 1/2, 0\}$ be such that $p_{a,t} \leq p_{b,t} \leq p_{c,t}$. If one lets $p'_{1,t}, p'_{1/2,t}, p'_{0,t}$ be the probabilities that voter t, with unknown urn j, will, respectively, vote for i, for an urn other than i and j, and for urn j, then we have that $p'_{1,t} \leq \frac{6}{5}p_{1,t}, p'_{1/2,t} \leq \frac{6}{5}p_{1/2,t}$ and $p'_{0,t} \leq \frac{6}{5}p_{0,t}$. Therefore $p'_{a,t} + p'_{b,t} \leq \frac{6}{5}(p_{a,t} + p_{b,t})$. It follows that $s'_t = 1 - p'_{c,t} = p'_{a,t} + p'_{b,t} \leq \frac{6}{5}(p_{a,t} + p_{b,t}) = \frac{6}{5}s_t$, which is upper bounded by $\frac{6}{5} \cdot s_t \leq \frac{6}{5} \cdot \frac{2}{3} = \frac{4}{5}$.

Observe that the sum s' of the s'_t's, $s' = \sum_{t=1}^{m} s'_t$, is then at most $\frac{6}{5}$ times the sum s of the s_t's; that is, $s' \leq \frac{6}{5}s \leq \frac{4}{5}\log_5 \eta^{-1}$.

Let X'_t be the random variable that, if the unknown urn is j, has value 1 if the t-th voter votes for urn i, 0 if she votes for urn j, and $1/2$ otherwise; we have:

$$\Pr\left[\forall t, X'_t \text{ equals the value } x_t \text{ that maximizes } \Pr[X_t = x_t]\right] = \prod_{t=1}^{m}(1 - s'_t).$$

Using the same greedy algorithm as before, but moving mass $x = \min\left(\frac{4}{5} - s'_{t'}, s'_t\right)$ from couples of s'_t's such that $s'_{t'} < \frac{4}{5}$ and $s'_t > 0$, $s'_{t'} \neq s'_t$, we get that the previous product is minimized when exactly $\left\lceil \frac{s'}{4/5} \right\rceil$ distinct s'_t's exist having value $4/5$, one having value $0 \leq s' - \left\lceil \frac{s}{4/5} \right\rceil \cdot \frac{4}{5} < \frac{4}{5}$, and the rest having null value. Then,

$$\begin{aligned}
\Pr\left[\forall t, X'_t \text{ equals the value } x_t \text{ that maximizes } \Pr[X_t = x_t]\right] &= \prod_{t=1}^{m}(1 - s'_t) \\
&\geq 5^{-\left\lceil \frac{s'}{4/5} \right\rceil - 1} \\
&\geq 5^{-\left\lceil \frac{4}{5} \cdot \frac{\log_5 \eta^{-1}}{4/5} \right\rceil - 1} \\
&\geq 5^{-\lceil \log_5 \eta^{-1} \rceil - 1} \geq \frac{\eta}{25}.
\end{aligned}$$

Now, if urn i won with this sequence of votes, it follows that j cannot win.

We have shown that if $\mathrm{Var}[X] \leq \frac{1}{72}\log_5 \eta^{-1}$, then the probability of winning is at most $1 - \frac{\eta}{25}$. We now assume $\mathrm{Var}[X] > \frac{1}{72}\log_5 \eta^{-1}$. We will use the following anti-concentration inequality (see Theorem 7.3.1 in [14], and [9]) to finish the proof:

Theorem 4.3 ([9, 14]). *Let $X = \sum_{i=1}^{n} X_i$, where X_i are independent random variables, with $X_i \in [0, 1]$, for $i = 1, \ldots, n$. Let $\sigma^2 = \mathrm{Var}[X]$ be $\sigma^2 \geq 40000$. Then, for each $t \in \left[0, \frac{\sigma^2}{100}\right]$, it holds that*

$$\Pr\left[X \geq E[X] + t\right] \geq c \cdot \exp\left(-\frac{t^2}{3\sigma^2}\right),$$

for some universal constant $c > 0$.

We apply Theorem 4.3 on the random variable $X = X(i, j)$, choosing $t = \sqrt{\frac{64e^7 m \,\mathrm{Var}[X]}{n^3\epsilon^{-2}}}$, if $\Delta_i(j) \geq 0$, and $t = 0$ otherwise. This choice is valid since

$$0 \leq \frac{t}{\mathrm{Var}[X]} \leq \sqrt{m \cdot \frac{64e^7}{n^3\epsilon^{-2}\,\mathrm{Var}[X]}} < \sqrt{m \cdot \frac{4608e^7 \ln 5}{n^3\epsilon^{-2}\ln \eta^{-1}}} \leq \frac{1}{100},$$

where the latter holds if $m \leq \frac{1}{46080000e^7 \ln 5} \cdot n^3\epsilon^{-2}\ln \eta^{-1}$.

We also need $\mathrm{Var}[X] \geq 40000$ to apply Theorem 4.3. Since $\mathrm{Var}[X] > \frac{1}{72}\log_5 \eta^{-1}$, and $\eta \leq H$, we choose H to be $H = 5^{-2880000}$, obtaining $\mathrm{Var}[X] > 40000$.

Observe that $E[X] = \frac{m}{2} - \frac{m}{2} \cdot \Delta_i(j)$. We show that the event "$X \geq E[X] + t$" implies the event "$X \geq \frac{m}{2}$" (which directly implies that the unknown urn i will not win the election).

If $\Delta_i(j) < 0$, the claim is trivial, since then $E[X] > \frac{m}{2}$, and t is non-negative. Otherwise, by the bound $\mathrm{Var}[X] \geq m \cdot \frac{|R_i - B_i|}{64}$, we get

$$t \geq m \cdot e^{7/2}\epsilon \cdot \sqrt{\frac{|R_i - B_i|}{n^3}} \geq \frac{m}{2} \cdot \Delta_i(j),$$

which proves that $X \geq E[X] + t \implies X \geq \frac{m}{2}$.

Applying Theorem 4.3, we get

$$\begin{aligned}
\Pr\left[X \geq E[X] + t\right] &\geq c \cdot \exp\left(-\frac{t^2}{3\,\mathrm{Var}[X]}\right) \\
&= c \cdot \exp\left(-\frac{64e^7 m}{3n^3\epsilon^{-2}}\right) \\
&\geq c \cdot \exp\left(-\frac{64}{3 \cdot 46080000 \cdot \ln 5} \cdot \ln \eta^{-1}\right) \\
&\geq c \cdot \eta.
\end{aligned}$$

The proof is then complete. □

5 An Upper Bound for Many Signals

In this section we consider the voting problem in its full generality: we have a set of $n \geq 2$ urns, with each urn $i = 1, \ldots, n$ inducing a distinct probability distribution $P_i = (p_{i,1}, p_{i,2}, \ldots, p_{i,C})$ over a set of C signals or colors. Let ϵ be the minimum ℓ_1 distance between the distributions P_i:

$$\epsilon = \min_{i \neq j} \ell_1(P_i, P_j) = \min_{i \neq j} \sum_{c=1}^{C} \left|p_{i,c} - p_{j,c}\right|.$$

Observe that when $C = 2$, this parameter ϵ is twice the one that we used in Section 2.

Theorem 5.1. *There exists a voting scheme that, using $m = \Theta\left(\frac{(C\log C)^2 n^3}{\epsilon^2} \ln \frac{n}{\eta}\right)$ voters, guarantee that the unknown urn wins with probability at least $1 - \eta$.*

The proof of this Theorem spans two subsections (Sections 5.1 and 5.2). In Section 5.1 we generalize the bichromatic voting scheme so (a) to treat urns that are not "well-separated" as if they were the same – this virtually increases the separation parameter ϵ – and (b) to guarantee, under some conditions, that the equivalent of the M parameter of the bichromatic voting scheme of Section 2 is not just upper bounded by $O(n^2\epsilon^{-1})$, but is actually asymptotic to $\Theta(n^2\epsilon^{-1})$.

In Section 5.2, we use both these properties to devise a new voting scheme that uses the generalized bichromatic one as a black box. The main idea of the multicolor voting scheme is to force voters to view the urns as bichromatic ones: each voter will choose a color c at random, and consider each urn as a bichromatic urn with colors $c, \bar{c}$ — that is, she will imagine that there are only two colors: "c" and "any color other than c". Using this trick directly with the bichromatic voting scheme of section 2 would decrease to 0 the

minimum distance between urns in the worst case. We do not want the separation between urns to decrease — since that would increase the minimum number of voters needed for the election to be successful — this is where property (a) of the generalized bichromatic voting scheme becomes pivotal. Also, we need a way to aggregate the votes given to each single urn, in each of the $(c, \bar{c})$ bichromatic instances; this has to be done in a way that guarantees that the right urn will win with high probability. We manage to do this by leveraging on property (b).

5.1 A More Flexible Upper Bound with Two Signals

To build a framework that can be used to handle the case of $C > 2$ signals, it is useful to consider a more general formulation of the bichromatic problem in which certain options can induce identical distributions over signals (and hence be indistinguishable from each other). We present the analysis in the language of urns and colored balls.

Thus, suppose we have a collection of n urns, labeled p_i for $i = 1, 2, \ldots, n$. With a slight abuse of notation we let p_i and $1 - p_i$ be, respectively, the fraction of blue balls, and of red balls, in urn p_i. We assume w.l.o.g. that $0 \leq p_1 \leq p_2 \leq \cdots \leq p_n \leq 1$.

We assume that one urn is adversarially chosen as the correct one (we will also refer to this as the *unknown* urn). Then each player draws a ball from the urn and votes for the name of an urn based on the color they observe. For this general version with indistinguishable urns, we will be interested in the probability that the urn receiving the most votes has the same distribution as the correct one; this general formulation is for the sake of the multi-color case later.

We describe the strategy that the players will use to randomly choose which vote to cast. First of all, for some $n' \geq 10$, choose $0 \leq p'_1 < p'_2 < \cdots < p'_{n'} \leq 1$. Let $\epsilon = \min_{1 \leq i \leq n'-1} \left(p'_{i+1} - p'_i\right)$. We require that (a) for $1 \leq k \leq \left\lceil \frac{n'-1}{3} \right\rceil = K$ it holds that $p'_{k+1} - p'_k \leq 2\epsilon$ and $p'_{n'-k+1} - p'_{n'-k} \leq 2\epsilon$, and (b) $p'_{K+1} \leq (2K+1)\epsilon$ and $p'_{n'-K} \geq 1 - (2K+1)\epsilon$.

The p'_i's are called the *landmarks* of the voting scheme.

1. Let $b_k = \sum_{\ell=1}^{k-1} \frac{2 - (p'_{\ell+1} + p'_\ell)}{p'_{\ell+1} - p'_\ell}$ and $r_k = \sum_{\ell=k}^{n'-1} \frac{p'_{\ell+1} + p'_\ell}{p'_{\ell+1} - p'_\ell}$ for $k = 1, \ldots, n'$.

2. Let $\phi : \{p_1, \ldots, p_n\} \to \{1, \ldots, n'\}$ be a mapping from urns to landmarks' indices, defined so that $\phi(p_i) = k$ if k maximizes $p_i b_k + (1 - p_i) r_k$ (ties can be broken arbitrarily).

3. Then define $R = \sum_{k=1}^{n'} \left(\left(\left| \phi^{-1}(k) \right| + 1 \right) \cdot r_k \right)$ and $B = \sum_{k=1}^{n'} \left(\left(\left| \phi^{-1}(k) \right| + 1 \right) \cdot b_k \right)$, and set $M = \max(R, B)$.

4. The probability that a voter will vote for p_j if a blue ball is drawn is

$$\Pr_{\mathbf{P} \sim f(\text{blue})}[\mathbf{P} = p_j] = M^{-1} \cdot \left(b_{\phi(p_j)} + \frac{M - B}{n} \right) = B_j.$$

5. The probability that a voter will vote for p_j if a red ball is drawn is

$$\Pr_{\mathbf{P} \sim f(\text{red})}[\mathbf{P} = p_j] = M^{-1} \cdot \left(r_{\phi(p_j)} + \frac{M - R}{n} \right) = R_j.$$

It is easy to check that the two probability distributions $(B_1, B_2, \ldots, B_n)$ and $(R_1, R_2, \ldots, R_n)$ are well-defined (their values are non-negative and they both sum up to one). Observe that $B_i = B_j$ and $R_i = R_j$ if $\phi(p_i) = \phi(p_j)$.

For a given urn p_i, let k_i^+ be the smallest positive index such that $p'_{k_i^+} \geq p_i$, if such an index exists, and k_i^- be the largest index such that $p'_{k_i^-} \leq p_i$, again if the index exists; observe that at least one of k_i^+ and k_i^- has to exist since $n' \geq 10$. We show the following lemma:

Lemma 5.2. *For each $i = 1, \ldots, n$, $\phi(p_i)$ is either equal to k_i^+ or to k_i^-. If, for some i, we have $p_i = p'_k$ it follows that $\phi(p_i) = k_i^+ = k_i^- = k$.*

Proof. For an arbitrary k it holds that

$$\begin{aligned} p_i \cdot b_k + (1 - p_i) \cdot r_k &= p_i \cdot \sum_{\ell=1}^{k-1} \frac{2 - (p'_{\ell+1} + p'_\ell)}{p'_{\ell+1} - p'_\ell} + (1 - p_i) \cdot \sum_{\ell=k}^{n'-1} \frac{p'_{\ell+1} + p'_\ell}{p'_{\ell+1} - p'_\ell} \\ &= \sum_{\ell=1}^{k-1} \frac{2p_i}{p'_{\ell+1} - p'_\ell} + \sum_{\ell=k}^{n'-1} \frac{p'_{\ell+1} + p'_\ell}{p'_{\ell+1} - p'_\ell} - p_i \cdot \sum_{\ell=1}^{n'-1} \frac{p'_{\ell+1} + p'_\ell}{p'_{\ell+1} - p'_\ell}. \end{aligned}$$

The latter sum does not depend on k. Therefore $p_i b_k + (1 - p_i) r_k$ is maximized with a k that maximizes the total of the former two sums.

Suppose that k_i^- exists and that $k < k_i^-$ maximizes the expression; then, by increasing k to $k + 1 \leq k_i^-$, we would remove the term $\frac{p'_{k+1} + p'_k}{p'_{k+1} - p'_k}$ from the second sum, and add the term $\frac{2p_i}{p'_{k+1} - p'_k}$ to the first one. Since, by definition $p'_k < p'_{k+1} \leq p'_{k^-} \leq p_i$, we have $p'_{k+1} + p'_k < 2p_i$, and therefore the total value of the first two sums would increase. It follows that $k < k_i^-$ cannot maximize the expression.

Analogously, suppose then that k_i^+ exists and that $k > k_i^+$ maximizes the expression; by decreasing k to $k - 1 \geq k_i^+$, we remove the term $\frac{2p_i}{p'_k - p'_{k-1}}$ from the first sum, and add the term $\frac{p'_k + p'_{k-1}}{p'_k - p'_{k-1}}$ to the second one. Since $p_i \leq p'_{k^+} \leq p'_{k-1} < p'_k$, we obtain that $2p_i < p'_{k-1} + p'_k$ and therefore we increase the total value of the first two sums; thus $k > k_i^+$ does not maximize the expression. □

We now turn to computing the probability that a player will vote for p_j given that the correct, adversarially chosen, distribution is p_i:

$$\Pr_{\substack{\mathbf{X} \sim p_i \\ \mathbf{P} \sim f(\mathbf{X})}} [\mathbf{P} = p_j] = p_i \Pr_{\mathbf{P} \sim f(\text{blue})} [\mathbf{P} = p_j] + (1 - p_i) \Pr_{\mathbf{P} \sim f(\text{red})} [\mathbf{P} = p_j] = p_i B_j + (1 - p_i) R_j = E_i(j).$$

We compute the difference between the probabilities that a vote for urn p_i and a vote for urn p_j are cast, given that the correct urn is p_i:

$$\Delta_i(j) = E_i(i) - E_i(j).$$

We will lower-bound $\Delta_i(j)$ to bound the number of voters needed to let the voting scheme be successful with high probability.

Lemma 5.3. *For each $1 \leq i, j \leq n$, it holds that*

$$\Delta_i(j) \geq \begin{cases} \frac{|\phi(p_i) - \phi(p_j)|}{M} & \text{if } p_i = p'_{k_i^+} = p'_{k_i^-} \\ \frac{\max(|\phi(p_i) - \phi(p_j)| - 1, 0)}{M} & \text{otherwise} \end{cases}$$

Proof. We make the expression of $\Delta_i(j)$ explicit:

$$\Delta_i(j) = p_i\left(B_i - B_j\right) + (1-p_i)\left(R_i - R_j\right) = p_i \cdot \frac{b_{\phi(p_i)} - b_{\phi(p_j)}}{M} + (1-p_i)\cdot\frac{r_{\phi(p_i)} - r_{\phi(p_j)}}{M}.$$

Suppose first that $\phi(p_j) < k_i^-$; then

$$\begin{aligned} M\cdot\Delta_i(j) &= p_i \cdot \sum_{\ell=\phi(p_j)}^{\phi(p_i)-1} \frac{2-(p'_{\ell+1}+p'_\ell)}{p'_{\ell+1}-p'_\ell} - (1-p_i)\cdot\sum_{\ell=\phi(p_j)}^{\phi(p_i)-1}\frac{p'_{\ell+1}+p'_\ell}{p'_{\ell+1}-p'_\ell} = \sum_{\ell=\phi(p_j)}^{\phi(p_i)-1}\frac{2\cdot p_i-(p'_{\ell+1}+p'_\ell)}{p'_{\ell+1}-p'_\ell} \\ &\geq \sum_{\ell=\phi(p_j)}^{k_i^- -1}\frac{2\cdot p_i-(p'_{\ell+1}+p'_\ell)}{p'_{\ell+1}-p'_\ell} \geq \sum_{\ell=\phi(p_j)}^{k_i^- -1}\frac{2\cdot p'_{\ell+1}-(p'_{\ell+1}+p'_\ell)}{p'_{\ell+1}-p'_\ell} = \sum_{\ell=\phi(p_j)}^{k_i^- -1}\frac{p'_{\ell+1}-p'_\ell}{p'_{\ell+1}-p'_\ell} = \left|k_i^- - \phi(p_j)\right|, \end{aligned}$$

where the first inequality follows from $\phi(p_i) \geq k_i^-$ and the second from $p_i \geq p'_{k_i^-} \geq p'_{\ell+1}$.

If, instead, $\phi(p_j) > k_i^+$ we have:

$$\begin{aligned} M\cdot\Delta_i(j) &= -p_i \cdot \sum_{\ell=\phi(p_i)}^{\phi(p_j)-1} \frac{2-(p'_{\ell+1}+p'_\ell)}{p'_{\ell+1}-p'_\ell} + (1-p_i)\cdot\sum_{\ell=\phi(p_i)}^{\phi(p_j)-1}\frac{p'_{\ell+1}+p'_\ell}{p'_{\ell+1}-p'_\ell} = \sum_{\ell=\phi(p_i)}^{\phi(p_j)-1}\frac{(p'_{\ell+1}+p'_\ell)-2p_i}{p'_{\ell+1}-p'_\ell} \\ &\geq \sum_{\ell=k_i^+}^{\phi(p_j)-1}\frac{(p'_{\ell+1}+p'_\ell)-2p_i}{p'_{\ell+1}-p'_\ell} \geq \sum_{\ell=k_i^+}^{\phi(p_j)-1}\frac{(p'_{\ell+1}+p'_\ell)-2p'_\ell}{p'_{\ell+1}-p'_\ell} = \sum_{\ell=k_i^+}^{\phi(p_j)-1}\frac{p'_{\ell+1}-p'_\ell}{p'_{\ell+1}-p'_\ell} = \left|k_i^+ - \phi(p_j)\right|, \end{aligned}$$

where the first inequality follows from $\phi(p_i) \leq k_i^+$ and the second from $p_i \leq p'_{k_i^+} \leq p'_\ell$.

Also, recall that $\phi(p_i)$ is an index that maximizes $p_i b_{\phi(p_i)} + (1-p_i)r_{\phi(p_i)}$. Since

$$\Delta_i(j) = M^{-1}\cdot\left(\left(p_i b_{\phi(p_i)} + (1-p_i)r_{\phi(p_i)}\right) - \left(p_i b_{\phi(p_j)} + (1-p_i)r_{\phi(p_j)}\right)\right),$$

we have that $\Delta_i(j) \geq 0$ for each ordered couple of urns i, j. The statement follows. □

We now give an upper bound on the probability that the correct urn will be chosen by a voter. Note, somewhat counter-intuitively, that the probability of a correct vote is higher when this upper bound is smaller — this is because the $\Delta_i(j)$ are additive gaps, not multiplicative ones, and so by making the upper bound on the expected number of votes for the correct urn smaller, the gap $\Delta_i(j)$ becomes larger relative to the mean.

Lemma 5.4. *It holds that*

$$E_i(i) \leq \frac{2(n'-1)}{\epsilon M} + \frac{1}{n}.$$

Proof. Recall that the correct urn is p_i. We upper-bound the probability that a vote will actually go to p_i:

$$\begin{aligned} E_i(i) &= p_i \cdot M^{-1}\cdot\left(b_{\phi(p_i)} + \frac{M-B}{n}\right) + (1-p_i)\cdot M^{-1}\cdot\left(r_{\phi(p_i)} + \frac{M-R}{n}\right) \\ &\leq p_i\cdot\left(M^{-1}\cdot b_{\phi(p_i)} + \frac{1}{n}\right) + (1-p_i)\cdot\left(M^{-1}\cdot r_{\phi(p_i)} + \frac{1}{n}\right). \end{aligned}$$

Observe that, by the definition of ϵ, we have that $b_i = \sum_{\ell=1}^{i-1} \frac{2-(p'_{\ell+1}+p'_\ell)}{p'_{\ell+1}-p'_\ell}$ is at most $b_i \leq \frac{2(i-1)}{\epsilon}$, and that $r_i = \sum_{\ell=i}^{n'-1} \frac{p'_{\ell+1}+p'_\ell}{p'_{\ell+1}-p'_\ell} \leq \frac{2(n'-i)}{\epsilon}$. Thus,

$$E_i(i) \leq p_i \cdot \left(\frac{2(i-1)}{\epsilon M} + \frac{1}{n}\right) + (1-p_i) \cdot \left(\frac{2(n'-i)}{\epsilon M} + \frac{1}{n}\right) \leq \frac{2(n'-1)}{\epsilon M} + \frac{1}{n}.$$

□

We now give an upper bound on M. This will allow us to apply a Chernoff bound and prove the main theorem.

Lemma 5.5. *It holds that*

$$\frac{1}{81} \cdot \frac{(n'-1)\cdot(n+n')}{\epsilon} \leq M \leq 2 \cdot \frac{(n'-1)\cdot(n+n')}{\epsilon}.$$

Proof. Recall that $M = \max(R, B)$; we will upper bound $R+B$ to get an upper bound on M:

$$\begin{aligned} R+B &= \sum_{k=1}^{n'} \left(\left(\left|\phi^{-1}\left(p'_k\right)\right|+1\right)\cdot(r_k+b_k)\right) \leq \sum_{k=1}^{n'} \left(\left(\left|\phi^{-1}\left(p'_k\right)\right|+1\right)\cdot(r_1+b_{n'})\right) \\ &= \sum_{k=1}^{n'} \left(\left(\left|\phi^{-1}\left(p'_k\right)\right|+1\right)\cdot \sum_{\ell=1}^{n'-1} \frac{2}{p_{\ell+1}-p_\ell}\right) \\ &\leq \sum_{k=1}^{n'} \left(\left(\left|\phi^{-1}\left(p'_k\right)\right|+1\right)\cdot(n'-1)\cdot\frac{2}{\epsilon}\right) = \frac{2\cdot(n'-1)\cdot(n+n')}{\epsilon}. \end{aligned}$$

It follows that

$$M \leq \frac{2\cdot(n'-1)\cdot(n+n')}{\epsilon}. \tag{6}$$

We now move on to the lower bound. Recall that $n' \geq 10$, and that, for each $k = 1, 2, \ldots, \left\lceil \frac{n'-1}{3} \right\rceil = K$ (resp., $k = n'-K, \ldots, n'-1$) we have $p'_{k+1} - p'_k \leq 2\epsilon$; also, $p'_{K+1} \leq (2K+1)\epsilon$ and $p'_{n'-K} \geq 1-(2K+1)\epsilon$. Observe that

$$2K+1 = 2\left\lceil \frac{n'-1}{3} \right\rceil + 1 \leq 2\cdot\frac{n'+1}{3} + 1 \leq 2\cdot\frac{n'+\frac{5}{2}}{3} \leq 2\cdot\frac{n'+\frac{n'}{4}}{3} \leq \frac{5}{6}\cdot n'.$$

Since $\epsilon = \min_{1\leq i\leq n'-1} p'_{i+1} - p'_i$, we have $\epsilon \leq \frac{1}{n'-1}$, and

$$(2K+1)\epsilon \leq \frac{5}{6}\cdot\frac{n'}{n'-1} \leq \frac{25}{27}.$$

We are now ready to lower bound M:

$$\begin{aligned}
M &= \max(R,B) \geq \frac{R+B}{2} = \frac{1}{2}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot(r_k+b_k)\right)\\
&= \frac{1}{2}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\sum_{\ell=k}^{n'-1}\frac{p'_{\ell+1}+p'_\ell}{p'_{\ell+1}-p'_\ell}\right)+\frac{1}{2}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\sum_{\ell=1}^{k-1}\frac{2-(p'_{\ell+1}+p'_\ell)}{p'_{\ell+1}-p'_\ell}\right)\\
&\geq \frac{1}{2}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\sum_{\ell=\max(k,n'-K)}^{n'-1}\frac{p'_{\ell+1}+p'_\ell}{p'_{\ell+1}-p'_\ell}\right)+\\
&\quad\frac{1}{2}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\sum_{\ell=1}^{\min(k-1,K)}\frac{2-(p'_{\ell+1}+p'_\ell)}{p'_{\ell+1}-p'_\ell}\right)\\
&\geq \frac{1}{2}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\sum_{\ell=\max(k,n'-K)}^{n'-1}\frac{2(1-(2K+1)\epsilon)}{2\epsilon}\right)+\\
&\quad\frac{1}{2}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\sum_{\ell=1}^{\min(k-1,K)}\frac{2-2\cdot(2K+1)\epsilon}{2\epsilon}\right),
\end{aligned}$$

using the upper bound we previously obtained for $(2K+1)\epsilon$, we obtain:

$$\begin{aligned}
M &\geq \frac{1}{2}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\sum_{\ell=\max(k,n'-K)}^{n'-1}\frac{2}{27\epsilon}\right)+\frac{1}{2}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\sum_{\ell=1}^{\min(k-1,K)}\frac{2}{27\epsilon}\right)\\
&\geq \frac{1}{27\epsilon}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\left(n'-\max(k,n'-K)+\min(k-1,K)\right)\right)\\
&\geq \frac{1}{27\epsilon}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot\left(\min(n'-k,K)+\min(k-1,K)\right)\right),
\end{aligned}$$

observe that $(n'-k)+(k-1)=n'-1$ and therefore at least one of $(n'-k)$ and $k-1$ is at least

$$\frac{n'-1}{2}\geq\frac{\frac{3}{2}n'-\frac{3}{2}}{3}=\frac{n'+\frac{n'-3}{2}}{3}\geq\frac{n'+7/2}{3}>\left\lceil\frac{n'-1}{3}\right\rceil=K.$$

Therefore, at least one of $\min(n'-k,K),\min(k-1,K)$ is at least K. Then,

$$M\geq\frac{1}{27\epsilon}\cdot\sum_{k=1}^{n'}\left(\left(\left|\phi^{-1}(k)\right|+1\right)\cdot K\right)=\frac{K}{27\epsilon}\cdot\sum_{k=1}^{n'}\left(\left|\phi^{-1}(k)\right|+1\right)=\frac{K}{27\epsilon}\cdot\left(n+n'\right)\geq\frac{(n'-1)\cdot(n+n')}{81\cdot\epsilon}.$$

□

Therefore, going back to the probability that an urn identical to the correct urn is voted for, we have

$$E_i(i)\leq\frac{2(n'-1)}{\epsilon M}+\frac{1}{n}\leq\frac{2(n'-1)}{\epsilon\cdot\frac{1}{81}\cdot\frac{(n'-1)(n+n')}{\epsilon}}=\frac{162}{n+n'}$$

5.2 An Upper Bound for Many Signals

In this section we consider the voting problem in its full generality: we have a set of $n \geq 2$ urns, with each urn $i = 1, \ldots, n$ inducing a distinct probability distribution $P_i = (p_{i,1}, p_{i,2}, \ldots, p_{i,C})$ over a set of C colors. Let ϵ be the minimum ℓ_1 distance between the distributions P_i:

$$\epsilon = \min_{i \neq j} \ell_1(P_i, P_j) = \min_{i \neq j} \sum_{c=1}^{C} |p_{i,c} - p_{j,c}| \,.$$

It turns out that the bichromatic scheme from Section 5.1 has already laid much of the groundwork for the multi-color case. Each voter u will behave as follows:

1. First, u will choose a color $c = c(u)$ uniformly at random from among all the colors. Voter u will then imagine the urns as inducing a bichromatic instance by imagining all colors other than c as a single color $\bar{c}$. In this way, urn i becomes a bichromatic urn with distribution $(p_{i,c}, 1 - p_{i,c})$ over its two colors.

2. Then, voter u will choose an integer $t = t(u) \in \{0, 1, \ldots, T\}$ with $T = \lceil \log_3 C \rceil + 1$, in such a way that $\Pr[t = i] = \alpha^{-1} \cdot 3^{-T+i}$, where $\alpha = \sum_{i=0}^{T} 3^{-i}$.

   Observe that $\alpha < \sum_{i=0}^{\infty} 3^{-i} = \frac{3}{2}$.

3. Voter u will then apply the bichromatic voting scheme from Section 5.1 to choose which urn to vote for. She will set $\{p_1, p_2, \ldots, p_n\} = \{p_{1,c}, p_{2,c}, \ldots, p_{n,c}\}$, for $i = 1, \ldots, n$; the sequence of the p_i''s will be defined as follows:

   - first she will pick a subsequence according to the following *marking algorithm*: set $w_1 = 0$ and mark all the p_j's such that $p_j \leq 3^{-t} \cdot \epsilon$; if some unmarked p_j remains, let $i = 2$, and
   - set w_i to be the smallest unmarked p_j,
   - mark all the p_j's for which $|p_j - w_i| < 3^{-t} \cdot \epsilon$;
   - if some unmarked p_j remains, repeat; otherwise, if $w_i \neq 1$, set $w_{i+1} = 1$; then, stop.

4. let i^* be the length of the sequence $\{w_i\}$; if $i^* < 10$, the voter will add $10 - i^*$ elements to $\{w_i\}$: let i be such that $w_{i+1} - w_i$ is maximized; the voter will insert the values the values $w_i + \frac{1}{9}(w_{i+1} - w_i), w_i + \frac{2}{9}(w_{i+1} - w_i), \ldots, w_i + \frac{8}{9}(w_{i+1} - w_i)$ in the list, keeping it sorted.

   The size of the sequence $\{w_i\}$ will then be at least 10.

   The voter will then define the sequences $x_{i,1}, x_{i,2}$ as $x_{i,1} = \frac{2w_i + w_{i+1}}{3}$, $x_{i,2} = \frac{w_i + 2w_{i+1}}{3}$, for $i = 1, \ldots, i^* - 1$;

5. the voter then merges the sequences $w_i, x_{i,1}, x_{i,2}$, and sorts the resulting sequence increasingly; let $y_1 < y_2 < \cdots < y_{3i^*-2}$ be this sequence, and $\epsilon_{c,t}$ be its separation parameter: $\epsilon_{c,t} = \min_{i=1,\ldots,3i^*-3} y_{i+1} - y_i$.

6. then the voter adds elements to $\{y_i\}$ in such a way that:

   (a) the minimum separation between adjacent elements remains at least $\epsilon_{c,t}$,

   (b) if the list has length n', then for each $i = 1, \ldots, \left\lceil \frac{n'}{3} \right\rceil$, it holds that $y_{i+1} - y_i \leq 2\epsilon_{c,t}$, and

   (c) if the list has length n', then for each $i = \left\lfloor \frac{2n'}{3} \right\rfloor, \ldots, n' - 1$, it holds that $y_{i+1} - y_i \leq 2\epsilon_{c,t}$.

To do so, she applies the following algorithm:

6.1 if (b) is not satisfied, i.e., if the list $\{y_i\}$ has currently length n' and i is a minimal index i for which there exists elements y_i, y_{i+1} such that $i \leq \left\lceil \frac{n'}{3} \right\rceil$ and $y_{i+1} - y_i > 2\epsilon_{c,t}$, then insert a new element between y_i and y_{i+1} of value $y_i + \epsilon_{c,t}$; this will increase the length of the list; repeat this step as long as (b) is not satisfied;

6.2 if (c) is not satisfied, i.e., if the list $\{y_i\}$ has currently length n' and there is a maximal index i for which there exists elements y_i, y_{i+1} such that $i \geq \left\lfloor \frac{2n'}{3} \right\rfloor$ and $y_{i+1} - y_i > 2\epsilon_{c,t}$, then insert a new element between y_i and y_{i+1} of value $y_{i+1} - \epsilon_{c,t}$; repeat this step as long as (c) is not satisfied.

It is easy to prove that the above algorithm guarantees properties (a), (b) and (c). Let $n'_{c,t}$ be the length of the final sequence $\{y_i\}$, $K_{c,t} = \left\lceil \frac{n'_{c,t}}{3} \right\rceil$, and observe that the algorithm also guarantees that (d) $n'_{c,t} \geq 10$, (e) $n'_{c,t} \leq 3i^* - 2 + 2(K_{c,t} + 1) \leq 9i^* + 2 \leq 9n + 2$ and (f) $y_{K_{c,t}+1} \leq (2K_{c,t} + 1)\epsilon_{c,t}$ and $y_{n'_{c,t}-K_{c,t}} \geq 1 - (2K_{c,t} + 1)\epsilon_{c,t}$.

The just-defined bichromatic instance depends only on the original multi-colored instance, on c and on t — we use (c,t)-instance to refer to the bichromatic instance induced by c and t.

Observe that the separation parameter $\epsilon_{c,t}$ of the (c,t)-instance will be at least $\epsilon_{c,t} \geq 3^{-t-3} \cdot \epsilon$, since the w_i's are at distance of at least $3^{-t-2} \cdot \epsilon$ from each other[3], and $x_{i,1}$, $x_{i,2}$ split the interval between w_i and w_{i+1} in three equal parts — the subsequently added y_i's do not induce gaps smaller that $\epsilon_{c,t}$. Furthermore the number of landmarks of the (c,t)-instance will be $10 \leq n'_{c,t} \leq 9n + 2 \leq 10n$, since $n \geq 2$.

We are now ready to prove Theorem 5.1, using the machinery built in Section 5.1.

Proof of Theorem 5.1. First, given two urns P_i, P_j, we say that a color c is *useful for* i, j if $|p_{i,c} - p_{j,c}| > \frac{\epsilon}{3C}$. Observe that if $C_{i,j}$ is the set of useful colors for urns P_i, P_j, we have

$$\sum_{c \in C_{i,j}} |p_{i,c} - p_{j,c}| > \frac{2}{3} \cdot \epsilon.$$

Indeed, since there are only C colors, the contribution to the ℓ_1 distance between P_i and P_j of their non-useful colors is less than $C \cdot \frac{\epsilon}{3C} = \frac{\epsilon}{3}$. Given that the total distance is at least ϵ, it follows that the useful colors contribute by more than $\frac{2\epsilon}{3}$ to the ℓ_1 distance of P_i, P_j.

Suppose that i is the unknown urn. Let $p_i = p_{i,c}$ and $p_j = p_{j,c}$ in the (c,t)-bichromatic instance, for some c, t. Let $E_i^{(c,t)}(j)$ be the expected number of votes that a voter will give to urn j, if i is the unknown urn, in the (c,t)-bichromatic instance. The analysis of the bichromatic case, guarantees that

$$E_i^{(c,t)}(j) \leq E_i^{(c,t)}(i) \leq \frac{162}{n + n'_{c,t}} \leq \frac{162}{n}.$$

and that the difference $\Delta_i^{(c,t)}(j) = E_i^{(c,t)}(i) - E_i^{(c,t)}(j)$ is at least

$$\Delta_i^{(c,t)}(j) \geq \frac{\max\left(|\phi(p_i) - \phi(p_j)| - 1, 0\right)}{M_{c,t}}.$$

Fix a color $c \in C_{i,j}$ and let $t_{c,i,j}$ be the smallest non-negative integer such that

$$|p_{i,c} - p_{j,c}| \geq \epsilon \cdot 3^{-t_{c,i,j}}.$$

[3]Before step 4 they were at distance at least $3^{-t}\epsilon$ from each other, and step 4 could have added new w_i's at distance at least 3^{-t-2} from each other.

Since c is a useful color we have $|p_{i,c} - p_{j,c}| > \frac{\epsilon}{3C}$, and therefore $0 \leq t_{c,i,j} \leq \lceil \log_3 C \rceil + 1 = T$. By $\epsilon_{c,t} \geq 3^{-t-3} \cdot \epsilon$, we obtain

$$\epsilon_{c,t_{c,i,j}} > \frac{1}{81} |p_{i,c} - p_{j,c}| .$$

Since $|p_i - p_j| \geq \epsilon \cdot 3^{-t_{c,i,j}}$, the marking algorithm run by the voters will mark p_i and p_j at different iterations — therefore, there are at least three landmarks between p_i and p_j. It follows that $|\phi(p_i) - \phi(p_j)| \geq 2$. Therefore,

$$\Delta_i^{(c,t_{c,i,j})}(j) \geq \frac{1}{M_{c,t_{c,i,j}}} \geq \frac{1}{2 \cdot \frac{(n'_{c,t_{c,i,j}} - 1) \cdot (n + n'_{c,t_{c,i,j}})}{\epsilon_{c,t_{c,i,j}}}} \geq \frac{\epsilon_{c,t_{c,i,j}}}{22n(n'_{c,t_{c,i,j}} - 1)} \geq \frac{|p_{i,c} - p_{j,c}|}{17820 \cdot n^2}.$$

Then,

$$\begin{aligned}\Delta_i(j) &= \sum_{c=1}^{C} \sum_{t=1}^{T} \left(\frac{1}{C} \cdot \frac{1}{T} \cdot \Delta_i^{(c,t)}(j) \right) \geq \frac{1}{C \cdot T} \cdot \sum_{c \in C_{i,j}} \max_{t=1,\ldots,T} \Delta_i^{(c,t)}(j) \geq \frac{1}{C \cdot T} \cdot \sum_{c \in C_{i,j}} \Delta_i^{(c,t_{c,i,j})}(j) \\ &\geq \frac{1}{C \cdot T} \cdot \sum_{c \in C_{i,j}} \frac{|p_{i,c} - p_{j,c}|}{17820 \cdot n^2} \geq \frac{\epsilon}{26730 \cdot C \cdot T \cdot n^2} = x.\end{aligned}$$

and

$$E_i(j) = \sum_{c=1}^{C} \sum_{t=1}^{T} \left(\frac{1}{C} \cdot \frac{1}{T} \cdot E_i^{(c,t)}(j) \right) \leq \frac{162}{n}$$

We choose $m = \left\lceil 7 \cdot 10^{12} \cdot \frac{C^2 \cdot T^2 \cdot n^3}{\epsilon^2} \ln \frac{n}{\eta} \right\rceil = \Theta\left(\frac{(C \log C)^2 \cdot n^3}{\epsilon^2} \ln \frac{n}{\eta} \right)$ as the number of voters, and we apply Chernoff bound (see Theorem 2.2), on V_j: the number of votes to urn j in the election, if i is the unknown urn. Observe that $E[V_j] = m \cdot E_i(j)$, and furthermore:

$$\begin{aligned}\Pr\left[|V_j - E[V_j]| > \frac{x}{3} \cdot m \right] &= \Pr\left[|V_j - E[V_j]| > \frac{x}{3E_i(j)} \cdot E_i(j) \cdot m \right] \\ &\leq 2 \exp\left(-\frac{x^2}{27E_i^2(j)} \cdot E_i(j) \cdot m \right) \\ &\leq 2 \exp\left(-\frac{x^2}{27E_i(j)} \cdot m \right) \\ &\leq 2 \exp\left(-\frac{\epsilon^2}{27 \cdot 26730^2 \cdot C^2 \cdot T^2 \cdot n^4 \cdot \frac{162}{n}} \cdot m \right) \\ &\leq 2 \exp\left(-\frac{\epsilon^2}{27 \cdot 26730^2 \cdot C^2 \cdot T^2 \cdot n^4 \cdot \frac{162}{n}} \cdot m \right) \leq \frac{\eta}{n}.\end{aligned}$$

Applying the Union Bound over all the urns, we have that each single urn j will deviate by at most $\frac{x}{3}m$ from its expected number of votes with probability at least $1 - \eta$, and since the expected difference of the number of votes of urn i and j if i is the unknown urn is at least $m \cdot \Delta_i(j) \geq m \cdot x$, we have that urn i will win the election with probability at least $1 - \eta$. □

6 Other Voting Systems

In this section we study other important voting systems, assuming that there are two types of signals; that is, assuming bichromatic urns.

6.1 Cumulative Voting

We show that *cumulative voting* requires a smaller number of voters for the election to succeed with high probability. In fact, *cumulative voting* can be exploited to work with a number of voters as small as the number of samples used by the optimal centralized algorithm (that is, the algorithm that, after sampling the minimum number of balls, produces the right guess with high probability).

Like plurality voting, in the cumulative voting election system, each voter has a single vote to cast; unlike plurality voting, though, the voter can split her vote arbitrarily between the candidates:

Definition 6.1 (Cumulative Voting)**.** *Each voter assigns a score to each candidate, in such a way that no score is negative and the sum of the scores assigned by a voter is* 1*. The total score of a candidate is the sum of the scores assigned to that candidate by the voters. If there exists a candidate i having a total score larger than the total score of each other candidate $j \neq i$, then i is the winner of the election.*

Given urns $p_1, p_2, \ldots, p_n$, the voting scheme we propose for cumulative voting is directly derived from the plurality voting scheme we proposed earlier; in the new scheme, there are only two possible votes: if a voter picks a red (resp., blue) ball then she will vote $(R_1, R_2, \ldots, R_n)$ (resp., $(B_1, B_2, \ldots, B_n)$) — that is, she will assign a weight of R_i (resp., B_i) to candidate P_i, for $i = 1, 2, \ldots, n$. The R_i's and the B_i's are those that we defined for the plurality voting scheme.

Theorem 6.2. *Let urns $p_1, p_2, \ldots, p_n$ be given, with urn p_i having a p_i fraction of blue balls, and a $1 - p_i$ fraction of red balls. Let $p_1 < p_2 < \cdots < p_n$. Also, let ϵ be $\epsilon = \min_{1 \leq i \leq n-1}(p_{i+1} - p_i)$. Then, for Cumulative Voting, $O\left(\epsilon^{-2} \ln \frac{1}{\eta}\right)$ voters are sufficient to guarantee a probability of at least $1 - \eta$ that the correct urn wins the election.*

Proof. Choose some $\eta \in (0, 1)$, and suppose the number of players is

$$m = \left\lceil 150 \cdot \epsilon^{-2} \cdot \ln \frac{2}{\eta} \right\rceil.$$

We start by using the Chernoff bound to show that the number of voters that pick a ball of some color is concentrated. If $p_i \geq \frac{1}{2}$, let X be the number of voters picking a blue ball; otherwise let X be the number of voters picking a red ball. In both cases, X is the sum of iid binary random variables X_j, with $X_j = 1$ with probability $\max(p_i, 1 - p_i)$ and $X_j = 0$ with probability $\min(p_i, 1 - p_i)$. Then, $E[X] = m \cdot \max(p_i, 1 - p_i)$ and $\frac{m}{2} \leq E[X] \leq m$. By Chernoff bound,

$$\begin{aligned}\Pr\left[|X - E[X]| > \frac{\epsilon}{5} \cdot m\right] &\leq \Pr\left[|X - E[X]| > \frac{\epsilon}{5} \cdot E[X]\right] \leq 2 \cdot \exp\left(-\frac{\epsilon^2}{75} \cdot E[X]\right) \\ &\leq 2 \cdot \exp\left(-\frac{\epsilon^2}{75} \cdot \frac{m}{2}\right) \leq 2 \cdot \exp\left(\ln \frac{2}{\eta}\right) = \eta,\end{aligned}$$

that is, with probability $1 - \eta$ the absolute difference between the number m_b of blue (resp., the number m_r of red) balls picked and the expectation $p_i \cdot m$ ($(1 - p_i) \cdot m$) by at most $\frac{\epsilon}{5} m$.

For any $i, j \in [n]$, $i \neq j$, let $V_i(j)$ be the fractional number of votes to P_j in the random election, with unknown urn P_i. Let $D_i(j) = V_i(i) - V_i(j)$. Observe that urn i beats urn j in the election (with unknown urn i) iff $D_i(j) > 0$. The random variable $D_i(j)$ is the sum of m iid random variables $D'_i(j)$ each taking value $B_i - B_j$ if the corresponding voter picked a blue ball and $R_i - R_j$ if she picked a red ball; we now bound the span S of values of $D'_i(j)$ — that is, we bound $S = |(B_i - B_j) - (R_i - R_j)|$. Observe that $B_i - B_j = \frac{b_i - b_j}{M}$ and $R_i - R_j = \frac{r_i - r_j}{M}$. Also,

$$b_i - b_j = \sum_{\ell=0}^{i-1} \frac{2 - (p_{\ell+1} + p_\ell)}{p_{\ell+1} - p_\ell} - \sum_{\ell=0}^{j-1} \frac{2 - (p_{\ell+1} + p_\ell)}{p_{\ell+1} - p_\ell},$$

and

$$r_i - r_j = \sum_{\ell=i}^{n-1} \frac{p_{\ell+1} + p_\ell}{p_{\ell+1} - p_\ell} - \sum_{\ell=j}^{n-1} \frac{p_{\ell+1} + p_\ell}{p_{\ell+1} - p_\ell}.$$

Therefore $i \geq j$ iff $b_i - b_j \geq 0$ and $r_i - r_j \leq 0$ — which implies that $|(b_i - b_j) - (r_i - r_j)| = |b_i - b_j| + |r_i - r_j|$ and thus the span S of $D'_i(j)$ is equal to $S = |B_i - B_j| + |R_i - R_j|$. Furthermore,

$$|b_i - b_j| = \sum_{\ell=\min(i,j)}^{\max(i,j)-1} \frac{2 - (p_{\ell+1} + p_\ell)}{p_{\ell+1} - p_\ell} \leq \frac{2|i-j|}{\epsilon}.$$

Analogously, $|r_i - r_j| \leq \frac{2|i-j|}{\epsilon}$. It follows that the span of $D'_i(j)$ can be upper bounded by

$$S = |B_i - B_j| + |R_i - R_j| \leq \frac{4|i-j|}{\epsilon M}.$$

Observe that $D_i(j)$ is a linear function of the number m_b of blue balls picked (and the number $m_r = m - m_b$ of red balls picked):

$$D_i(j) = m_b \cdot (B_i - B_j) + m_r \cdot (R_i - R_j).$$

Therefore,

$$E[D_i(j)] = m \cdot p_i \cdot (B_i - B_j) + m \cdot (1 - p_i) \cdot (R_i - R_j) = m \cdot \Delta_i(j),$$

where $\Delta_i(j)$ is the functional defined in Section 2; recall that we proved there that $\Delta_i(j) \geq \frac{|i-j|}{M}$.

Recall that with probability $1 - \eta$, $|m_b - m \cdot p_i| \leq \frac{\epsilon}{5} \cdot m$. If this event happens we have that, for each $j \neq i$,

$$D_i(j) \geq E[D_i(j)] - \frac{\epsilon}{5} \cdot m \cdot S \geq m \cdot \left(\Delta_i(j) - \frac{4}{5} \cdot \frac{|i-j|}{M} \right) = m \cdot \frac{|i-j|}{5M} > 0.$$

Therefore, urn i will beat each urn $j \neq i$, with probability $1 - \eta$. The proof is concluded. □

Observe that the previous bound is tight in a strong sense: no algorithm that picks $o\left(\epsilon^{-2} \ln \frac{1}{\eta}\right)$ balls, and produces a guess arbitrarily after having seen all their colors, is able to guess the right urn with probability at least $1 - \eta$.

6.2 Condorcet Voting

In this section we present a conjecture, and we elaborate on it, with the aim of showing that *Condorcet voting* is as good as Cumulative voting — and is thus optimal. We begin by recalling the definition of the Condorcet voting system:

Definition 6.3 (Condorcet Voting)**.** *In a Condorcet election, each voter returns a (total) ordering of the candidates. Given two candidates i and j, we say that i beats j in a run-off election if more than half the voters ranked i higher than j. If there exists a candidate i that beats each other candidate $j \neq i$ in a run-off election, then i is the winner of the Condorcet election.*

We observe that, in Condorcet voting, voters do not assign real numbers to candidates as in Cumulative voting — they rather return a discrete object: a permutation of them.

There exist many variants of the Condorcet election. The differences between them lie in the way of dealing with ties (that is, when no candidate i beats each other candidate j in a run-off election). Our main theorem holds for any such variant, since our theorem will guarantee that no ties will exist with high probability.

We start by defining a set of coefficients that will be useful for introducing a Condorcet voting scheme.

Definition 6.4. *For $k \geq 0$ and $\ell \geq 0$, let $c_{k,\ell}$ be:*

$$c_{k,\ell} = (-1)^{k+\ell} \cdot \binom{k+\ell}{k} - \frac{(-1)^{\ell} \cdot \binom{k+\ell}{k} + (-1)^{k+\ell} \cdot (\ell+1)}{k+\ell+2}.$$

We define inductively the sequence $b_{k,\ell}$, indexed by two integers; if $k < 0$ or $\ell < 0$ then $b_{k,\ell} = 0$; for notational simplicity we will use:

$$x_{k,\ell} = \sum_{\substack{i,j \geq 0 \\ i+j \leq k}} \frac{b_{i,k-i-j} \cdot b_{j,\ell}}{(i+1)(i+j+2)}, \qquad y_{k,\ell} = \sum_{i=0}^{k} \left((-1)^{k-i} \cdot \frac{b_{i,\ell}}{i+1} \right).$$

Then, for $k \geq 0, \ell \geq 0$, we define $b_{k,\ell} = (k+1) \cdot (x_{k-1,\ell} + y_{k-1,\ell} - c_{k+1,\ell})$.

The induction is well-defined: $x_{k-1,\ell}$ only depends on the $b_{i,j}$'s for which either (a) $i+j \leq k-1$, or (b) $i \leq k-1$ and $j = \ell$; $y_{k-1,\ell}$ only depends on the $b_{i,j}$'s for which $i \leq k-1$ and $j = \ell$. Therefore, the $b_{k,\ell}$'s can be computed in the following order via the recurrence: for $n = 0, 1, 2, \ldots$ and for $k = 0, \ldots, n$, compute $b_{k,n-k}$. We now make a conjecture on the $b_{k,\ell}$'s:

Conjecture 6.5. *Let $B(x,P) = \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \left(b_{k,\ell} \cdot x^k \cdot P^{\ell} \right)$. Then,*

(i) the series $B(x,P)$ converges for $0 \leq x < \frac{1}{2}$, $x \leq P \leq \frac{1}{2}$,

(ii) $B(x,P) \geq 0$ for $0 \leq x < \frac{1}{2}$, $x \leq P \leq \frac{1}{2}$, and

(iii) $\int_0^P B(x,P)\,dx = \frac{1}{P+1} - \sqrt{\frac{1-2P}{1+2P}}$, for $0 \leq P \leq \frac{1}{2}$.

Comments on Conjecture 6.5. We now make some comments on the conjecture, indicating some possible approaches to settle it.

(i) Numerical approximations indicate that $B(x,P)$ converges for each $0 \leq x < \frac{1}{2}$, $x \leq P \leq \frac{1}{2}$, but it diverges for $x = P = \frac{1}{2}$.

(ii) For $k \geq 0$, and $0 \leq \ell \leq k$, let

$$a_{k,\ell} = \sum_{i=0}^{\ell} \left(\binom{k+2}{i} \cdot b_{k,\ell-i} \right),$$

also, let

$$B_1(x,P) = \sum_{k=0}^{\infty} \left(\frac{\sum_{\ell=0}^{k} \left(a_{k,\ell} \cdot P^{\ell} \right)}{(P+1)^{k+2}} \cdot x^k \right).$$

By looking at the first few terms of $B_1(x,P)$'s Taylor expansion, it appears that $B(x,P) = B_1(x,P)$.

The $B_1(x,P)$ expression, if $B_1(x,P) = B(x,P)$, could be quite useful to prove non-negativity, since the $a_{k,\ell}$'s seem all to be non-negative — if they are point (ii) of the conjecture directly follows.

Also, if $B_1(x,P) = B(x,P)$, then one can express $b_{k,\ell}$ (for each $\ell \geq 0$) in terms of $b_{k,0}, b_{k,1}, \ldots, b_{k,k}$:

$$b_{k,\ell} = \sum_{i=0}^{k} \left(b_{k,i} \cdot \sum_{j=i}^{\min(k,\ell)} \left(\frac{(-1)^{\ell-j}}{(\ell-j)!} \cdot \binom{k+2}{j-i} \cdot \sum_{h=0}^{\ell-j} \left(\begin{bmatrix} \ell-j \\ h \end{bmatrix} \cdot (k+2)^h \right) \right) \right),$$

where $\begin{bmatrix} n \\ k \end{bmatrix}$ represents the unsigned Stirling number of the first kind with indices $n \geq k$. The last claim can be proved using the $B_1(x,P)$ expression and the following expression for $(Q+1)^{-t}$, $t > 0$:

$$\frac{1}{(Q+1)^t} = \sum_{i=0}^{\infty} \left(\frac{(-1)^i}{i!} \cdot \sum_{j=0}^{i} \left(\begin{bmatrix} i \\ j \end{bmatrix} \cdot t^j \right) \cdot Q^i \right).$$

$b_{k,\ell}$	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$
$k=0$	1	-2	3	-4	5	-6
$k=1$	2	-2	0	4	-10	18
$k=2$	3	-8	18	-36	65	-108
$k=3$	$\frac{20}{3}$	$-\frac{44}{3}$	20	$-\frac{44}{3}$	$-\frac{40}{3}$	80
$k=4$	$\frac{25}{3}$	$-\frac{64}{3}$	53	$-\frac{388}{3}$	$\frac{880}{3}$	-610
$k=5$	$\frac{98}{5}$	$-\frac{844}{15}$	$\frac{582}{5}$	-188	$\frac{668}{3}$	$-\frac{558}{5}$

Table 1: The first few $b_{k,\ell}$'s.

$a_{k,\ell}$	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$	$\ell=5$
$k=0$	1					
$k=1$	2	4				
$k=2$	3	4	4			
$k=3$	$\frac{20}{3}$	$\frac{56}{3}$	$\frac{40}{3}$	$\frac{16}{3}$		
$k=4$	$\frac{25}{3}$	$\frac{86}{3}$	50	$\frac{106}{3}$	$\frac{32}{3}$	
$k=5$	$\frac{98}{5}$	$\frac{1214}{15}$	$\frac{2012}{15}$	$\frac{656}{5}$	$\frac{1016}{15}$	$\frac{46}{3}$

Table 2: The first few $a_{k,\ell}$'s.

(iii) Since

$$\frac{1}{P+1}-\sqrt{\frac{1-2P}{1+2P}}=\sum_{n=0}^{\infty}\left(C_n\cdot P^n\right),$$

with

$$C_n=\begin{cases}1-\binom{2\lfloor n/2\rfloor}{\lfloor n/2\rfloor} & \text{if } n \text{ is even}\\ 2\binom{2\lfloor n/2\rfloor}{\lfloor n/2\rfloor}-1 & \text{if } n \text{ is odd}\end{cases}=\frac{1-3\cdot(-1)^n}{2}\cdot\binom{2\,\lfloor n/2\rfloor}{\lfloor n/2\rfloor}+(-1)^n,$$

we have that the point (iii) of the conjecture states that, for $n\geq 0$,

$$\sum_{k=0}^{n}\frac{b_{k,n-k}}{k+1}=C_{n+1}.$$

Tables 1 and 2 show how the $b_{k,\ell}$ and the $a_{k,\ell}$ sequences begins.

□

We use Conjecture 6.5 to show the existence of a probability distribution over permutations that induces a given set of "marginals".

Lemma 6.6. *Let $0\leq p_1<p_2<\cdots<p_n\leq 1$. Then, if Conjecture 6.5 is true, there exists a probability distribution over the symmetric group S_n, such that, for each $1\leq i<j\leq n$,*

$$\Pi_{i,j}=\Pr_{\pi}\left[\pi(i)<\pi(j)\right]=\min\left(1,\frac{1}{p_i+p_j}\right).$$

Proof. To each probability $p\in[0,1]$ we associate a random variable X_p with values:

- if $p\leq\frac{1}{2}$, X_p is the constant random variable p, with a point mass $\gamma_p=1$ at p;

- if $p > \frac{1}{2}$, X_p has a point mass of $0 \leq \gamma_p \leq 1$ at p; X_p has a density function $f_p(x)$, with $f_p(x) = \alpha_p(x)$ for $x \in \left[1-p, \frac{1}{2}\right]$ and $f_p(x) = \beta_p(x)$ for $x \in \left(\frac{1}{2}, p\right)$, with

$$\alpha_p(x) = (p+x)^{-2},$$

$$\beta_p(x) = B\left(x - \frac{1}{2}, p - \frac{1}{2}\right),$$

and

$$\gamma_p = \sqrt{\frac{1}{p} - 1}.$$

By Conjecture 6.5, we have that $\beta_p(x) \geq 0$ and $\int_{1/2}^{p} \beta_p(x)\,dx = \frac{2}{2p+1} - \sqrt{\frac{1}{p} - 1}$; therefore the total probability mass assigned to X_p is $1 - X_p$ is then a well-defined random variable.

The CDF associated to $\alpha_p(x)$, for $p \geq \frac{1}{2}$ and $1 - p \leq y \leq \frac{1}{2}$ is:

$$\int_{1-p}^{y} \alpha_p(x)dx = \frac{y - 1 + p}{y + p}.$$

The CDF associated to $\beta_p(x)$, for $p \geq \frac{1}{2}$ and $\frac{1}{2} \leq y \leq p$ is:

$$\int_{\frac{1}{2}}^{y} \beta_p(x)dx = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty}\left(\frac{b_{k,\ell}}{k+1} \cdot \left(y - \frac{1}{2}\right)^{k+1} \cdot \left(p - \frac{1}{2}\right)^{\ell}\right)$$

Observe that if $p \neq q$, then there's zero probability that $X_p = X_q$ — since, for each p, X_p has a positive point mass only at p. Then, we pick a permutation π of $\{1, 2, \ldots, n\}$ by letting $\pi(i) < \pi(j)$ iff $X_{p_i} < X_{p_j}$.

We verify that the marginals $\Pi_{i,j}$ of our distribution satisfy the requirement in the claim:

- if $0 \leq p < q \leq 1 - p$, then:

$$\Pr\left[X_p \leq X_q\right] = 1$$

- if $0 \leq p < \frac{1}{2}$ and $1 - p \leq q \leq 1$, then:

$$\Pr\left[X_p \leq X_q\right] = \Pr\left[X_q \geq p\right] = 1 - \int_{1-q}^{p} (q+x)^{-2}\,dx = \frac{1}{p+q}.$$

- if $\frac{1}{2} < p < q \leq 1$. Let $P_1 = \Pr\left[X_p \leq X_q \leq \frac{1}{2}\right]$, $P_2 = \Pr\left[X_p \leq \frac{1}{2} \leq X_q\right]$, $P_3 = \Pr\left[\frac{1}{2} \leq X_p \leq p \leq X_q\right]$, and $P_4 = \Pr\left[\frac{1}{2} \leq X_p \leq X_q < p\right]$. Then,

$$\Pr\left[X_p \leq X_q\right] = P_1 + P_2 + P_3 + P_4.$$

We now show that $P_1 + P_2 + P_3 + P_4 = \frac{1}{p+q}$, completing the proof. We start by computing P_1 and P_2:

$$\begin{aligned} P_1 &= \int_{1-p}^{\frac{1}{2}} \left((q+x)^{-2} \cdot \int_{1-p}^{x} (p+y)^{-2}dy\right)dx = \int_{1-p}^{\frac{1}{2}} \frac{x - (1-p)}{(q+x)^2(p+x)}dx \\ &= \frac{2p-1}{(q-p)(2q+1)} - \frac{\ln\frac{(2p+1)(1-p+q)}{2q+1}}{(q-p)^2}. \end{aligned}$$

$$P_2 = \int_{1-p}^{\frac{1}{2}} (p+x)^{-2}dx \cdot \left(1 - \int_{1-q}^{\frac{1}{2}} (q+x)^{-2}dx\right) = \frac{2p-1}{2p+1} \cdot \left(1 - \frac{2q-1}{2q+1}\right) = \frac{2p-1}{2p+1} \cdot \frac{2}{2q+1}.$$

As for P_3 and P_4, we have:

$$P_3 = \left(\int_{\frac{1}{2}}^{p} \beta_p(x)dx + \sqrt{\frac{1}{p} - 1}\right) \cdot \left(\int_{p}^{q} \beta_q(x)dx + \sqrt{\frac{1}{q} - 1}\right) = \frac{2}{2p+1} \cdot \left(\frac{2}{2q+1} - \int_{\frac{1}{2}}^{p} \beta_q(x)dx\right).$$

$$P_4 = \int_{\frac{1}{2}}^{p} \left(\beta_q(x) \cdot \int_{\frac{1}{2}}^{x} \beta_p(y)dy\right) dx$$

We show that for our definition of $B(x, P) = \beta_{P+1/2}(x + 1/2)$, the equation $\Pr[X_p \leq X_q] = P_1 + P_2 + P_3 + P_4$ is equal to $P_1 + P_2 + P_3 + P_4 = \frac{1}{p+q}$, for each $\frac{1}{2} < p < q \leq 1$.

The latter is true iff $P_3 + P_4 = \frac{1}{p+q} - P_1 - P_2$; expanding the terms, we get the equation holds iff:

$$\frac{2}{2p+1} \cdot \left(\frac{2}{2q+1} - \int_{\frac{1}{2}}^{p} \beta_q(x)dx\right) + \int_{\frac{1}{2}}^{p} \left(\beta_q(x) \cdot \int_{\frac{1}{2}}^{x} \beta_p(y)dy\right) dx =$$
$$\frac{1}{p+q} - \frac{2p-1}{(q-p)(2q+1)} + \frac{\ln \frac{(q-p+1)(2p+1)}{2q+1}}{(q-p)^2} - \frac{2p-1}{2p+1} \cdot \frac{2}{2q+1},$$

or, equivalently,

$$\int_{\frac{1}{2}}^{p} \left(\beta_q(x) \cdot \int_{\frac{1}{2}}^{x} \beta_p(y)dy\right) dx - \frac{2}{2p+1} \cdot \int_{\frac{1}{2}}^{p} \beta_q(x)dx =$$
$$\frac{1}{p+q} + \frac{\ln \frac{(q-p+1)(2p+1)}{2q+1}}{(q-p)^2} - \frac{2q-1}{(q-p)(2q+1)}.$$

For notational convenience, we let $P = p - \frac{1}{2}$ and $Q = q - \frac{1}{2}$. Let $B(P, Q) = \beta_q(p) = \beta_{Q+\frac{1}{2}}(P + \frac{1}{2})$. We thus need to prove:

$$\int_{0}^{P} \left(B(x, Q) \cdot \int_{0}^{x} B(y, P)dy\right) dx - \frac{1}{P+1} \cdot \int_{0}^{P} B(x, Q)dx =$$
$$\frac{1}{P+Q+1} + \frac{\ln \frac{(Q-P+1)(P+1)}{Q+1}}{(Q-P)^2} - \frac{Q}{(Q+1)(Q-P)}. \tag{7}$$

Since we will expand the right-hand side in a Taylor series around $(Q, P) = (0, 0)$, we observe that the right-hand side has a removable singularity at $P = Q$. Indeed, letting $P = Q - \epsilon$,

$$\lim_{\epsilon \to 0} \left(\frac{\ln \frac{(1+\epsilon)(Q+1-\epsilon)}{Q+1}}{\epsilon^2} - \frac{Q}{(Q+1)\epsilon}\right) = \lim_{\epsilon \to 0} \left(\frac{\ln\left((1+\epsilon)\left(1 - \frac{\epsilon}{Q+1}\right)\right)}{\epsilon^2} - \frac{Q}{(Q+1)\epsilon}\right)$$
$$= \lim_{\epsilon \to 0} \left(\frac{\epsilon - \frac{\epsilon}{Q+1} - \frac{\epsilon^2}{Q+1} - \frac{Q^2\epsilon^2}{2(Q+1)^2} + O(\epsilon^3)}{\epsilon^2} - \frac{Q}{(Q+1)\epsilon}\right)$$
$$= \lim_{\epsilon \to 0} \frac{-\frac{\epsilon^2}{Q+1} - \frac{Q^2\epsilon^2}{2(Q+1)^2} + O(\epsilon^3)}{\epsilon^2}$$
$$= -\frac{1}{Q+1} - \frac{Q^2}{2(Q+1)^2},$$

since the limit from the left and the right coincide. Therefore, if $P = Q$, the right-hand side of Equation 7 is

$$\frac{1}{2Q+1} - \frac{1}{Q+1} - \frac{Q^2}{2(Q+1)^2}.$$

Recalling that

$$B(P,Q) = \sum_{\ell=0}^{\infty}\sum_{k=0}^{\infty}\left(b_{k,\ell}\cdot P^k\cdot Q^\ell\right),$$

we get that the following holds:

$$\begin{aligned}\int_0^P\left(B(x,Q)\cdot\int_0^x B(y,P)dy\right)dx &= \sum_{\ell=0}^{\infty}\sum_{k=2}^{\infty}\left[\left(\sum_{\substack{i,j\geq 0\\ i+j\leq k-2}}\frac{b_{i,k-2-i-j}\cdot b_{j,\ell}}{(i+1)(i+j+2)}\right)\cdot P^k\cdot Q^\ell\right]\\ &= \sum_{\ell=0}^{\infty}\sum_{k=2}^{\infty}\left(x_{k-2,\ell}\cdot P^k\cdot Q^\ell\right).\end{aligned}$$

Also,

$$\int_0^P B(x,Q)dx = \sum_{\ell=0}^{\infty}\sum_{k=1}^{\infty}\left(\frac{b_{k-1,\ell}}{k}\cdot P^k\cdot Q^\ell\right),$$

and, since $\frac{1}{P+1} = \sum_{k=0}^{\infty}\left((-1)^k\cdot P^k\right)$,

$$-\frac{1}{P+1}\cdot\int_0^P B(x,Q)dx = -\sum_{\ell=0}^{\infty}\sum_{k=1}^{\infty}\left[\sum_{i=0}^{k-1}\left((-1)^{k-1-i}\cdot\frac{b_{i,\ell}}{i+1}\right)\cdot P^k\cdot Q^\ell\right] = -\sum_{\ell=0}^{\infty}\sum_{k=1}^{\infty}\left(y_{k-1,\ell}\cdot P^k\cdot Q^\ell\right).$$

Furthermore, we have

$$\begin{aligned}\frac{1}{P+Q+1} &= \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty}c_1(k,\ell)\cdot P^k\cdot Q^\ell\\ \frac{\ln\frac{(Q-P+1)(P+1)}{Q+1}}{(Q-P)^2} - \frac{Q}{(Q+1)(Q-P)} &= \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty}c_2(k,\ell)\cdot P^k\cdot Q^\ell\end{aligned}$$

with

$$c_1(k,\ell) = (-1)^{k+\ell}\cdot\binom{k+\ell}{k},$$

$$c_2(k,\ell) = -\frac{(-1)^\ell\cdot\binom{k+\ell}{k} + (-1)^{k+\ell}\cdot(\ell+1)}{k+\ell+2}.$$

observe that $c_{k,\ell} = c_1(k,\ell) + c_2(k,\ell)$, where $c_{k,\ell}$ is as in Definition 6.4.

We then have that $P_1 + P_2 + P_3 + P_4 = \frac{1}{p+q}$ (particularly, Equation 7) is satisfied iff

$$\sum_{\ell=0}^{\infty}\sum_{k=0}^{\infty}\left((x_{k-2,\ell} - y_{k-1,\ell})\cdot P^k\cdot Q^\ell\right) = \sum_{\ell=0}^{\infty}\sum_{k=0}^{\infty}\left(c_{k,\ell}\cdot P^k\cdot Q^\ell\right).$$

The equality holds iff for each $k,\ell\geq 0$, we have

$$c_{k,\ell} = x_{k-2,\ell} - y_{k-1,\ell}.$$

Observe that if $k = 0$, then the equation is satisfied, since $c_{0,\ell} = x_{-2,\ell} = y_{-1,\ell} = 0$. If $k \geq 1$, then $y_{k-1,\ell} = \frac{b_{k-1,\ell}}{k} - y_{k-2,\ell}$. Since we defined the $b_{k,\ell}$'s to be, for for each $k, \ell \geq 0$,

$$b_{k,\ell} = (k+1) \cdot (x_{k-1,\ell} + y_{k-1,\ell} - c_{k+1,\ell}),$$

we have that Equation 7 is satisfied and $P_1 + P_2 + P_3 + P_4 = \frac{1}{p+q}$. □

Using the previous distribution over permutations we can prove the main theorem of the section.

Theorem 6.7. *Let $0 \leq p_1 < p_2 < \cdots < p_n \leq 1$ be the blue-probabilities of a set of bichromatic urns. Let $\epsilon = \min_{1 \leq i \leq n-1}(p_{i+1} - p_i)$. If Conjecture 6.5 is true, there exists a symmetric voting scheme for the Condorcet election that guarantees that the unknown urn wins with probability $1 - \eta$ with $O\left(\frac{\ln \eta^{-1}}{\epsilon^2}\right)$ voters.*

Proof. Lemma 6.6 guarantees the existence of a probability distribution P over the set of permutations of $\{1, 2, \ldots, n\}$ such that, for each $1 \leq i < j \leq n$, $\Pr_{\pi \sim P}[\pi(i) < \pi(j)] = \min\left(1, \frac{1}{p_i + p_j}\right)$. If $\pi(j) > \pi(i)$ we say that j beats i in π.

We also let $q_i = 1 - p_i$; therefore $0 \leq q_n < q_{n-1} < \cdots < q_1 \leq 1$, and $\min_{1 \leq i \leq n-1}(q_{i+1} - q_i) = \epsilon$. Lemma 6.6 again guarantees the existence of a probability distribution Q over the set of permutations of $\{1, 2, \ldots, n\}$ such that for $n \geq i > j \geq 1$, we have $\Pr_{\pi \sim Q}[\pi(i) < \pi(j)] = \min\left(1, \frac{1}{q_i + q_j}\right)$.

Each voter will apply the following algorithm: if she draws blue, she sample a permutation according to P, otherwise she samples a permutation according to Q.

Now, suppose the i-th urn is the unknown urn. Let $j \neq i$ be the index of any other urn. If $j > i$,

$$\begin{aligned}\Pr[\text{the unknown urn } i \text{ beats another urn } j] &= p_i \Pr_{\pi \sim P}[\pi(i) > \pi(j)] + (1 - p_i) \Pr_{\pi \sim Q}[\pi(i) > \pi(j)] \\ &= p_i \max\left(0, 1 - \frac{1}{p_i + p_j}\right) + q_i \min\left(1, \frac{1}{q_i + q_j}\right),\end{aligned}$$

if $p_i + p_j \leq 1$ (and therefore $q_i + q_j \geq 1$) the latter simplifies to $\frac{q_i}{q_i + q_j}$; otherwise $p_i + p_j > 1, q_i + q_j < 1$, and the expression simplifies to $p_i\left(1 - \frac{1}{p_i + p_j}\right) + q_i = p_i - \frac{p_i}{p_i + p_j} + 1 - p_i = 1 - \frac{p_i}{p_i + p_j} = \frac{p_j}{p_i + p_j}$.

Therefore, if $j > i$, we have

$$\Pr[\text{the unknown urn } i \text{ beats another urn } j] \in \left\{\frac{q_i}{q_i + q_j}, \frac{p_j}{p_i + p_j}\right\}.$$

If, otherwise, $j < i$, we have

$$\begin{aligned}\Pr[\text{the unknown urn } i \text{ beats another urn } j] &= p_i \Pr_{\pi \sim P}[\pi(i) > \pi(j)] + (1 - p_i) \Pr_{\pi \sim Q}[\pi(i) > \pi(j)] \\ &= p_i \min\left(1, \frac{1}{p_i + p_j}\right) + q_i \max\left(0, 1 - \frac{1}{q_i + q_j}\right),\end{aligned}$$

if $q_i + q_j \leq 1$ (and therefore $p_i + p_j \geq 1$) the latter simplifies to $\frac{p_i}{p_i + p_j}$; otherwise $q_i + q_j > 1, p_i + p_j < 1$, and the expression simplifies to $p_i + q_i\left(1 - \frac{1}{q_i + q_j}\right) = 1 - q_i + q_i - \frac{q_i}{q_i + q_j} = \frac{q_j}{q_i + q_j}$.

Therefore, in any case, $\Pr[\text{the unknown urn } i \text{ beats another urn } j] \in \left\{\frac{\max(q_i, q_j)}{q_i + q_j}, \frac{\max(p_i, p_j)}{p_i + p_j}\right\}$. We lower-bound the latter two fractions:

$$\begin{aligned}\frac{\max(p_i, p_j)}{p_i + p_j} &= \frac{\max(p_i, p_j)}{2\max(p_i, p_j) - |p_i - p_j|} = \frac{\max(p_i, p_j) - \frac{1}{2}|p_i - p_j|}{2\max(p_i, p_j) - |p_i - p_j|} + \frac{1}{2} \cdot \frac{|p_i - p_j|}{2\max(p_i, p_j) - |p_i - p_j|} \\ &= \frac{1}{2} \cdot \left(1 + \frac{|p_i - p_j|}{p_i + p_j}\right) \geq \frac{1}{2} + \frac{|p_i - p_j|}{4},\end{aligned}$$

and, analogously,

$$\frac{\max(q_i, q_j)}{q_i + q_j} \geq \frac{1}{2} + \frac{|q_i - q_j|}{4}.$$

Since, $|q_i - q_j| = |p_i - p_j|$, we have

$$\Pr[\text{the unknown urn } i \text{ beats another urn } j] \geq \frac{1}{2} + \frac{|p_i - p_j|}{4} \geq \frac{1}{2} + \frac{|i-j|}{4} \cdot \epsilon.$$

Now, given two urns i, j, let $X_i(j)$ be the random variable counting the number of votes in which $i > j$, with m voters. Observe that if $X_i(j) > \frac{m}{2}$, then urn i beats urn j. Also,

$$\frac{m}{2} \leq m \cdot \left(\frac{1}{2} + \frac{|i-j| \cdot \epsilon}{4}\right) \leq E[X_i(j)] \leq m,$$

and

$$\begin{aligned}
\Pr\left[|X_i(j) - E[X_i(j)]| \geq \frac{|i-j| \cdot \epsilon}{5} \cdot m\right] &\leq \Pr\left[|X_i(j) - E[X_i(j)]| \geq \frac{|i-j| \cdot \epsilon}{5} \cdot E[X_i(j)]\right] \\
&\leq \exp\left(-\frac{|i-j|^2 \cdot \epsilon^2}{75} \cdot E[X_i(j)]\right) \\
&\leq \exp\left(-\frac{|i-j|^2 \cdot \epsilon^2}{150} \cdot m\right).
\end{aligned}$$

By choosing $m = \left\lceil 150\epsilon^{-2} \ln \frac{3}{\eta} \right\rceil$, we obtain:

$$\Pr\left[|X_i(j) - E[X_i(j)]| \geq \frac{|i-j| \cdot \epsilon}{5} \cdot m\right] \leq \exp\left(-|i-j|^2 \ln \frac{3}{\eta}\right) = \left(\frac{\eta}{3}\right)^{|i-j|^2} \leq \left(\frac{\eta}{3}\right)^{|i-j|}.$$

Observe that if $|X_i(j) - E[X_i(j)]| < \frac{|i-j| \cdot \epsilon}{5} \cdot m$, then — by $E[X_i(j)] \geq m \cdot \left(\frac{1}{2} + \frac{|i-j|\epsilon}{4}\right)$ — we get $X_i(j) \geq m \cdot \left(\frac{1}{2} + \frac{|i-j|\epsilon}{20}\right) > \frac{m}{2}$, which implies that urn i beats urn j.

Applying the Union Bound over all the urns $j \neq i$, we obtain

$$\Pr[\text{urn } i \text{ does not win the election}] \leq 2 \cdot \sum_{k=1}^{\infty} \left(\frac{\eta}{3}\right)^k = 2 \cdot \frac{\eta/3}{1 - \eta/3} \leq \eta.$$

□

Acknowledgments. We thank Larry Blume, David Easley, and Bobby Kleinberg for valuable discussions.

References

[1] Lisa R. Anderson and Charles A. Holt. Information cascades in the laboratory. *American Economic Review*, 87(5):847–862, December 1997.

[2] David Austen-Smith and Jeffrey S. Banks. Information aggregation, rationality, and the Condorcet Jury Theorem. *American Political Science Review*, 90(1):34–45, March 1996.

[3] Marco Battaglini, Rebecca B. Morton, and Thomas R. Palfrey. The swing voter's curse in the laboratory. *Review of Economic Studies*, 77:61–89, 2010.

[4] Sourav Bhattacharya. Preference reversal and information aggregation in elections, November 2007. Working paper, U. Pitt Dept. of Economics.

[5] Ioannis Caragiannis and Ariel D. Procaccia. Voting almost maximizes social welfare despite limited communication. *Artificial Intelligence*, 175:1655–1671, 2011.

[6] Timothy J. Feddersen and Wolfgang Pesendorfer. The swing voter's curse. *American Economic Review*, 86(3):408–424, June 1996.

[7] Timothy J. Feddersen and Wolfgang Pesendorfer. Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *American Political Science Review*, 92(1):23–35, March 1998.

[8] Timothy J. Feddersen and Wolfgang Pesendorfer. Abstention in elections with asymmetric information and diverse preferences. *American Political Science Review*, 93(2):381–398, June 1999.

[9] W. Feller. Generalization of a probability limit theorem of cramér. *Transactions of the American Mathematical Society*, 54:361–372, 1943.

[10] Dino Gerardi and Leeat Yariv. Deliberative voting. *Journal of Economic Theory*, 134(1):317–338, May 2007.

[11] Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, March 2007.

[12] Patrick Hummel. Jury theorems with multiple alternatives. *Social Choice and Welfare*, 34(1):65–103, 2010.

[13] Patrick Hummel. Information aggregation in multicandidate elections under plurality rule and runoff voting. *Mathematical Social Sciences*, to appear.

[14] Jiří Matoušek and Jan Vondrák. The probabilistic method.

[15] Elchanan Mossel, Allan Sly, and Omer Tamuz. From agreement to asymptotic learning. Technical Report arXiv:1105.4765v3 [math.ST], arxiv.org, June 2011.

[16] Roger B. Myerson. Extended Poisson games and the Condorcet Jury Theorem. *Games and Economic Behavior*, 25(1):111–131, October 1998.

[17] H. Peyton Young. Condorcet's theory of voting. *American Political Science Review*, 82(4):1231–1244, December 1988.